

REPORT DOCUMENTATION PAGE			<i>Form Approved</i> <i>OMB No. 0704-0188</i>	
<small>Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.</small>				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE 22.Oct.98	3. REPORT TYPE AND DATES COVERED THESIS	
4. TITLE AND SUBTITLE WARM SEASON STATISTICAL VERIFICATION OF THE PENN. STATE. UNIV.REAL TIME MESOSCALE MODEL VERSION 5			5. FUNDING NUMBERS	
6. AUTHOR(S) CAPT FITZGERALD MARK P				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) PENNSLYVANIA STATE UNIVERSITY			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) THE DEPARTMENT OF THE AIR FORCE AFIT/CIA, BLDG 125 2950 P STREET WPAFB OH 45433			10. SPONSORING/MONITORING AGENCY REPORT NUMBER 98-094	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION AVAILABILITY STATEMENT Unlimited distribution In Accordance With AFI 35-205/AFIT Sup 1			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words)			19990106 047	
14. SUBJECT TERMS			15. NUMBER OF PAGES	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT			18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT
				20. LIMITATION OF ABSTRACT

Author: Mark Patrick Fitzgerald

Thesis Title: WARM-SEASON STATISTICAL VERIFICATION OF THE
PENNSYLVANIA STATE UNIVERSITY REAL-TIME
MESOSCALE MODEL VERSION 5

Rank/Branch: Captain/Air Force

Date: May 1998

Number of Pages: 48

Degree Awarded: Master of Science in Meteorology

Institution: The Pennsylvania State University

DTIC QUALITY ASSURANCE

ABSTRACT

For this study, warm-season simulations from the real-time Pennsylvania State University / National Center for Atmospheric Research (PSU/NCAR) mesoscale model, known widely as MM5, are verified against surface METAR and upper-air soundings. Verification statistics are calculated on two model domains. The 36-km coarse-mesh domain encompasses the CONUS and portions of the surrounding regions. The 12-km fine-mesh domain is centered over Pennsylvania and encompasses the surrounding states. Variables that are verified include temperature, dew point, relative humidity, wind direction, wind speed, geopotential height, sea-level pressure, as well as the total totals severe weather index. Verification statistics calculated include bias, mean absolute error, root mean square error, and the decomposition of a skill score based on the mean square error.

Coarse-mesh surface results show an early-morning warm and dry bias, a positive wind speed bias, and wind directions that veer clockwise with respect to the observations. A cool bias is present on the coarse mesh at 850 mb for all forecast times, possibly indicating an over-prediction of boundary layer depth. Interestingly, the fine mesh does not perform better than the coarse mesh overall. But, over the 24-hour forecast period, fine-mesh skill markedly improves and nearly matches coarse-mesh skill for many variables, especially the moisture fields. By improving the fine-mesh initialization procedure, the fine-mesh forecasts should prove more skillful than the coarse-mesh forecasts.

REFERENCES

- Anthes, R.A., 1983: Regional Models of the Atmosphere in Middle Latitudes. *Monthly Weather Review*, 111, 1306-1335.
- Blackadar, A. K., 1979: High-resolution models of the planetary boundary layer. *Advances in Environmental Science and Engineering*, 1, No. 1. Pfafflin and Ziegler, Eds., Gordon and Breach Sci. Publ., New York, 50-85.
- Brier, G. W., and R. A. Allen, 1951: Verification of Weather Forecasts. *Compendium of Meteorology*, T.F. Malone, Ed., American Meteorological Society, 841-848.
- Cox, R., B. L. Bauer, T. Smith, 1998: A Mesoscale Model Intercomparison. *Bulletin of the American Meteorological Society*, Vol. 79, No. 2, 265-283.
- Doswell, C. A., 1986: Short Range Forecasting. *Mesoscale Meteorology and Forecasting*, P. S. Ray, Ed., American Meteorological Society, 793 pp.
- Dudhia, J., 1993: A nonhydrostatic version of the Penn State / NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Monthly Weather Review*, 121, 1493-1513.
- Environmental Sciences Services Administration, 1968: Climatic Atlas of the United States. U. S. Department Of Commerce, U. S. Government Printing Office, Washington DC, 80pp.
- Fairall, L. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young, 1996: Bulk Parameterization of Air Sea Fluxes for TOGA COARE. *Journal of Geophysical*

Research, Vol. 101, 3734-3764.

Grell, G. A., J. Dudhia, D. R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). *NCAR Technical Note-398+STR*, National Center for Atmospheric Research (NCAR), Boulder, 138pp.

Kain, J. S. and J. M. Fritsch, 1993: Convective parameterization in mesoscale models: The Kain-Fritsch scheme. *The representation of cumulus convection in numerical models*. Meteorological Monographs, No. 46, American Meteorological Society. K. A. Emanuel, and D. J. Raymond, EDS. 165-170.

Manning, K. W., and C. A. Davis, 1997: Verification and Sensitivity Experiments for the WISP94 MM5 Forecasts. *Weather And Forecasting*, Vol. 12 No. 4, 719-735.

Murphy, A. H., 1995: Coefficients of Correlation and Determination as Measures of Performance in Forecast Verification. *Weather And Forecasting*, Vol. 10 No. 4, 681-688.

_____, A. H., 1988: Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. *Monthly Weather Review*, Vol. 116 2417-2424.

_____, A. H., and H. Daan, 1985: Forecast Evaluation. *Probability, Statistics, and Decision Making in the Atmospheric Sciences*, A. H. Murphy and R. W. Katz, Eds., Westview Press, 379-437.

_____, A. H., and R. L. Winkler, 1987: A General Framework for Forecast Verification. *Monthly Weather Review*, Vol. 115, 1330-1338.

Oncley, S. P., and J. Dudhia, 1995: Evaluation of Surface Fluxes from MM5 Using Observations. *Monthly Weather Review*, Vol. 123, 3344-3357.

Shafran, P. C., 1997: Summertime Circulation Patterns in the Lake Michigan Area. *MS Thesis, The Pennsylvania State University*, 206pp.

White, B. G., J. Paegle, W. J. Steenburgh, J. D. Horel, R. T. Swanson, D. J. Onton, and J. Miles, 1998: Short-Term Forecast Validation of Six Models for Winter 1996. *Weather and Forecasting*, in press.

Zhang, D. L. and R.A. Anthes, 1982: A high-resolution model of the planetary boundary layer-sensitivity tests and comparison with SESAME-79 data. *Journal of Applied Meteorology*, 114, 1330-1339.

The Pennsylvania State University

The Graduate School

Department of Meteorology

WARM-SEASON STATISTICAL VERIFICATION OF
THE PENNSYLVANIA STATE UNIVERSITY
REAL-TIME MESOSCALE MODEL VERSION 5

A Thesis in

Meteorology

by

Mark Patrick Fitzgerald

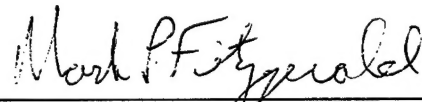
© 1998 Mark Patrick Fitzgerald

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

May 1998

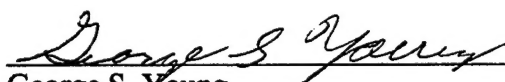
I grant The Pennsylvania State University the nonexclusive right to use this work for the University's own purposes and to make single copies of the work available to the public on a not-for-profit basis if copies are not otherwise available.

A handwritten signature in cursive script, reading "Mark P. Fitzgerald". The signature is written in dark ink and is positioned above a horizontal line.

Mark P. Fitzgerald

We approve the thesis of Mark P. Fitzgerald.

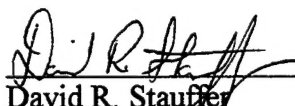
Date of Signature



George S. Young

Associate Professor of Meteorology and Geo Environmental Engineering
Thesis Advisor

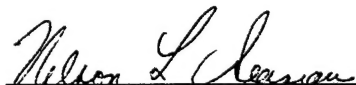
10 April, 1998



David R. Stauffer

Research Associate
Second Reader

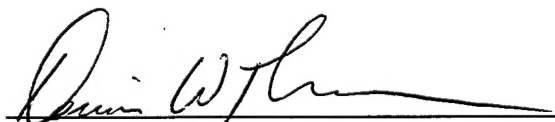
14 April 1998



Nelson L. Seaman

Assistant Professor of Meteorology
Third Reader

16 April 1998



Dennis W. Thomson

Professor of Meteorology
Head of the Department of Meteorology

10.4.98

ABSTRACT

For this study, warm-season simulations from the real-time Pennsylvania State University / National Center for Atmospheric Research (PSU/NCAR) mesoscale model, known widely as MM5, are verified against surface METAR and upper-air soundings. Verification statistics are calculated on two model domains. The 36-km coarse-mesh domain encompasses the CONUS and portions of the surrounding regions. The 12-km fine-mesh domain is centered over Pennsylvania and encompasses the surrounding states. Variables that are verified include temperature, dew point, relative humidity, wind direction, wind speed, geopotential height, sea-level pressure, as well as the total totals severe weather index. Verification statistics calculated include bias, mean absolute error, root mean square error, and the decomposition of a skill score based on the mean square error.

Coarse-mesh surface results show an early-morning warm and dry bias, a positive wind speed bias, and wind directions that veer clockwise with respect to the observations. These errors are consistent with systematic inaccuracies in the surface-layer fluxes of the Blackadar boundary-layer scheme. A cool bias is present on the coarse mesh at 850 mb for all forecast times, possibly indicating an over-prediction of boundary layer depth. Interestingly, the fine mesh does not perform better than the coarse mesh overall. But, over the 24-hour forecast period, fine-mesh skill markedly improves and nearly matches coarse-mesh skill for many variables, especially the moisture fields, by the end of the forecast. By improving the fine-mesh initialization procedure, the fine-mesh forecasts should become more skillful than the coarse-mesh forecasts.

TABLE OF CONTENTS

LIST OF FIGURES.....	V
ACKNOWLEDGMENTS	VI
CHAPTER 1 INTRODUCTION.....	1
1.1 MOTIVATION AND GOALS FOR THE VERIFICATION SYSTEM.....	1
1.2 THE PENNSYLVANIA STATE UNIVERSITY MESOSCALE MODELING SYSTEM.....	2
CHAPTER 2 SIMULATION AND OBSERVATION DESCRIPTION	4
2.1 MODEL SIMULATIONS.....	4
2.2 OBSERVATIONAL DATA	4
CHAPTER 3 STATISTICAL ANALYSIS METHODOLOGY.....	5
3.1 OVERVIEW	5
3.2 STANDARD VERIFICATION STATISTICS	5
3.3 MURPHY SKILL SCORE STATISTICS.....	6
CHAPTER 4 VERIFICATION RESULTS	9
4.1 COARSE-MESH FINDINGS	9
4.1.1 Temperature	9
4.1.2 Dew Point.....	12
4.1.3 Relative Humidity.....	14
4.1.4 Wind Direction	15
4.1.5 Wind Speed	16
4.1.6 Geopotential Height.....	17
4.1.7 Sea Level Pressure.....	18
4.2 FINE-MESH FINDINGS	18
4.2.1 Temperature	18
4.2.2 Dew Point.....	19
4.2.3 Relative Humidity.....	20
4.2.4 Wind Direction	21
4.2.5 Wind Speed	22
4.2.6 Geopotential Height.....	22
4.2.7 Sea Level Pressure.....	23
4.3 COARSE AND FINE-MESH MODEL VERTICAL STRUCTURE VERIFICATION	23
4.4 COARSE MESH VERSUS FINE MESH	24
CHAPTER 5 CONCLUSIONS	26
5.1 SUMMARY OF MM5 STRENGTHS AND WEAKNESSES.....	26
5.2 FUTURE WORK.....	29
REFERENCES.....	30
APPENDIX FIGURES.....	33

LIST OF FIGURES

Figure 1. Coarse-mesh temperature statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 2. Surface temperature bias, forecast time = t_0+12	Appendix
Figure 3. Coarse-mesh dew-point statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 4. Dew-point skill-score at 700 mb, forecast time = t_0+12	Appendix
Figure 5. Coarse-mesh relative-humidity statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 6. Coarse-mesh wind-direction statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 7. Coarse-mesh wind-speed statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 8. Coarse-mesh height statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 9. Coarse-mesh sea-level pressure statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 10. Sea-level pressure bias, forecast time = t_0+12	Appendix
Figure 11. Fine-mesh temperature statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 12. Surface temperature bias, forecast time = t_0+12	Appendix
Figure 13. Fine-mesh dew-point statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 14. Dew-point bias at 250 mb, forecast time = t_0+24	Appendix
Figure 15. Fine-mesh relative-humidity statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 16. Fine-mesh wind-direction statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 17. Fine-mesh wind-speed statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 18. Fine-mesh height statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 19. Fine-mesh sea-level pressure statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 20. Total-totals index statistics for 1 June 1997 to 28 August 1997	Appendix
Figure 21. Coarse-mesh and fine-mesh statistical comparison	Appendix

ACKNOWLEDGMENTS

I would like to thank Dr. George Young for leading me through this academic challenge. What helped me most is his ability to chart a clear path through tasks that I thought monumental, while prioritizing along the way. I also extend my gratitude to Dr. David Stauffer. Any time I needed guidance, Dr. Stauffer made himself available, gave me 100% of his attention, and guided me toward relevant papers and scientifically sound methods and solutions. Thanks go to Dr. Nelson Seaman for helping keep all of my work focused and innovative. Glenn Hunter was often the first person I would approach with a question. I knew he was willing to help and could always find an answer, usually wedged in between two jokes. He is a great part of the team here. Thanks Glenn. Thanks also go to Annette Lario-Gibbs. Annette knows the MM5 code by heart, all five gigs. She was always happy to get me answers and even write code modules when needed. Her great ability to communicate helped to quickly complete each complex task. My appreciation goes to Dr. Ray Masters for his many hours of help and guidance when it came time to visualize and put into figures the mountain of statistical results. For help with generating the final two-dimensional figures I thank Robert Hart. Thanks go to Daryl Onton and Kirby Cook from the University of Utah for their help with GEMPAK, an MM5-to-GEMPAK conversion, and verification concepts. For help with data handling, mesoscale meteorology, statistics, and atmospheric sounding concepts, thanks go to Art Person, Dr. J. Michael Fritsch, Dr. Bob Vislocky, and George Brian, respectively. Thanks go to the United States Air Force for sending me to a great university.

My wife Lynette deserves pages of thanks. Lynette, thanks for motivating me in the mornings, making killer lunches, and understanding that I had to spend so much time at school. You are the greatest person on earth! Thanks go to Mom, Peggy Larkin and Dad, Frank Fitzgerald. They always let me know I could do anything I put my heart into.

CHAPTER 1

INTRODUCTION

1.1 Motivation and Goals for the Verification System

The objective of this study is to identify the strengths and weaknesses of the version of the Pennsylvania State University / National Center for Atmospheric Research (PSU/NCAR) mesoscale model, known widely as MM5, operated in real time at Penn State. Quantifying its skill will enable the model developers to pinpoint problem areas within the model and will help guide them toward possible ways to improve its performance. For example, Manning and Davis (1997) used wintertime statistics to test specific MM5 parameterization changes. Moreover, the availability of documented verification information on the real-time MM5 products will allow forecasters and researchers to more intelligently use these products in various applications. This verification system is intended to provide the foundation for what will become an intense and ongoing exploration into the strengths, biases, and deficiencies of a greater variety of forecast variables, derived products, and parameterizations that are being generated by or utilized in, the real-time MM5 model at Penn State.

The statistics produced by this verification system may also be used to compare the Penn State real-time MM5 forecasts to the forecasts from other mesoscale models in use today. For example, White et al. (1998) statistically compare the MM5 to five other mesoscale models for the winter of 1996. Additionally, Cox et al. (1998) use statistics from a variety of climates and seasons for an inter-comparison of four mesoscale models,

one of which is the MM5. The University of Utah statistically verifies the MM5, National Weather Service MESO-ETA, and NCEP research models over each season and for special cases (Kirby L. Cook 1998, personal communication). The statistical results presented in this study are for the warm season from 1 June to 28 August 1997. Unlike many published statistical studies of mesoscale models, we present in detail the statistical values so that they can be used in future comparisons.

1.2 The Pennsylvania State University Mesoscale Modeling System

The real-time mesoscale model at Penn State is the non-hydrostatic Penn State/NCAR MM5 (Version 2) (Grell et al. 1994, Dudhia 1993). Forecasts verified here were produced on two grids. The outer domain covers the continental United States with a 36-km mesh, while the inner domain over the Northeastern U. S. has a 12-km mesh. The inner domain has finer-resolution terrain to better represent complex topography and dynamic processes. The model utilizes 30 terrain-following (sigma) vertical coordinate layers with approximately 9 of those layers below 1 km AGL. Key physical parameterizations in the present version include: Blackadar planetary boundary layer (Blackadar 1979), explicit moist physics with a simple ice phase (no mixed-phase processes), Kain-Fritsch deep convection (Kain and Fritsch 1993), and atmospheric radiation. For a more complete description of the PSU/NCAR MM5 model see Grell et al. (1994).

Preparation of initial and lateral boundary conditions begins with interpolation of the National Centers for Environmental Prediction's (NCEP) current "early-Eta" model

analyses and forecast fields to the MM5 outer mesh. Each 36-km run of the MM5 is started 12 h prior to the beginning of the forecast period ($t = t_0 - 12$ h). Over that 12-h period, Newtonian relaxation (Stauffer and Seaman 1990) is used in a dynamic initialization to gradually assimilate the analyses, so that the model's fields are in approximate dynamic balance at the beginning of the forecast period and condensed-phase processes are well developed. Finally, the 12-km fine grid, integrated for 3 h prior to the beginning of the forecast period at $t = t_0 - 3$ h, uses interpolated 36-km dynamically initialized fields for initial conditions and lateral boundary conditions. No data are assimilated on the fine-mesh through this 3-h pre-forecast period. The $t = t_0 + 00$ h corresponds to 0000 UTC.

Forecasts are run in parallel on a 4-processor SGI R10000 Power Challenge. Following the dynamic initialization, the optimized MM5 code produces a 36-h forecast on the 36-km outer domain and a 24-h forecast on the 12-km domain. Including the initialization, forecasts for both meshes are completed in less than five hours. The daily model output including convective indices is available on the web at (<http://mm5.met.psu.edu/mm5/>).

CHAPTER 2

SIMULATION AND OBSERVATION DESCRIPTION

2.1 Model Simulations

Model simulations are archived hourly on the native sigma-vertical-coordinate levels. A software post-processor is used to interpolate from the sigma levels to pressure levels. Model evaluation is performed at the model surface layer (i.e. lowest model sigma level, about 30 m AGL), as well as the 850, 700, 500, and 250-mb levels. The coarse mesh is verified at the $t = t_0+00$ h analysis time, $t = t_0+12$ h, $t = t_0+24$ h, and $t = t_0+36$ h forecast times. The fine mesh is verified at the $t = t_0+00$ h, $t = t_0+12$ h, and $t = t_0+24$ h times. Model variables that are verified include temperature, dew point, relative humidity, wind direction, wind speed, geopotential height, sea level pressure, and the total totals index, which is used as an aid in forecasting severe convective weather.

2.2 Observational Data

National Weather Service rawinsonde and METAR-surface observations are used for verification of the MM5 forecasts. To help prevent erroneous observations from being used in the verification, a quality-control check is done first on all temperature, wind, and sea-level pressure observations. To do this we compare each observation to the ETA analysis value for that time and level. If the observation does not equal the ETA value, plus or minus an assigned tolerance range, the observation is assumed to be unreliable and is thrown out. The tolerance range is based on Penn State's experience with real-data input into MM5. For most of the warm season cases studied here, between 1% and 6% of the observations were discarded.

CHAPTER 3

STATISTICAL ANALYSIS METHODOLOGY

3.1 Overview

The objective of this verification package is to illuminate the strengths and weaknesses of the MM5, and to infer which physical parameterizations are performing well or need improvement. Since no one verification score can give a complete understanding of the capability of a forecasting system (Brier and Allen 1951, Murphy and Daan 1985), we calculate a selection of scores designed to fulfill the verification study objectives.

3.2 Standard Verification Statistics

We calculate common verification measures such as the bias, root mean square error (RMSE), and mean absolute error (MAE), which are useful for comparing forecast systems (Murphy and Winkler 1987). The bias is a direct measure of the tendency of the model to over-forecast or under-forecast a particular variable. The RMSE, which is always positive, represents the error in the forecast in a way that emphasizes the magnitude of the extreme errors. The MAE is also always positive, but measures the average or typical magnitude of error in the model for a particular variable. Anthes (1983) gives a description of bias and RMSE and includes tables of typical values for both. Murphy and Daan (1985) discuss bias, RMSE, and MAE as well as many other skill scores. The bias, RMSE and MAE are defined as

$$Bias = \sum_N \left[\frac{(f - o)}{N} \right] \quad (1)$$

$$RMSE = \left\{ \sum_N \left[\frac{(f - o)^2}{N} \right] \right\}^{\frac{1}{2}} \quad (2)$$

$$MAE = \sum_N \left[\frac{|f - o|}{N} \right] \quad (3)$$

where f and o are the forecast and observation values respectively and N is the total number of forecast and observation pairs evaluated.

3.3 Murphy Skill Score Statistics

Bias, RMSE, and MAE measures are commonly used for comparing models, but to further explore strengths and weaknesses of, and ways to improve, a forecasting system, Murphy and Winkler (1987) demonstrate that other diagnostic measures can be valuable. For this purpose, they outline two factorizations of the joint distribution of forecasts and observations leading to new statistical scores. We have chosen to utilize a Skill Score (SS), equation (12) in Murphy (1988), that is based on the mean square error (MSE) and falls into the calibration-refinement factorization. Calibration refers to a pattern comparison between the observations and model forecast and refinement refers to model “willingness” to forecast values away from the mean (Murphy and Winkler 1987). The SS is the ‘sum’ of three non-negative parts and is defined as

$$SkillScore(SS) = COD - MULTE - ADDE \quad (4)$$

where,

$$COD = r^2 = \left[\frac{N \sum \frac{(f * o)}{N} - \left(\sum \frac{f}{N} \right) \left(\sum \frac{o}{N} \right)}{\sqrt{\left[N \sum \frac{f^2}{N} - \left(\sum \frac{f}{N} \right)^2 \right] \left[N \sum \frac{o^2}{N} - \left(\sum \frac{o}{N} \right)^2 \right]}} \right]^2 \quad (5)$$

$$MULTE = \left[r - (S_f / S_o) \right]^2 \quad (6)$$

$$ADDE = \left[\frac{(\bar{f} - \bar{o})}{S_o} \right]^2. \quad (7)$$

The term r in (5) is the coefficient of correlation, S_f is the standard deviation of the forecasts, and S_o is the standard deviation of the observations. The SS can have values ranging from 1, a perfect score, to $-\infty$, the worst possible score.

The COD, or the coefficient of determination, is the square of the coefficient of correlation, r , and is a measure of the linear relationship between the forecasts and observations. The COD is a measure of how well the forecast *pattern* matches the observed *pattern* of a particular variable and is not a measure of how closely the magnitudes match. The COD can be thought of as the skill score that is achieved if the model output statistics (MOS) (Glahn and Lowry 1972) technique is applied to a single predictor (the model forecast of the variable being evaluated). The forecast is perfectly correlated with the observations when the COD has a value of 1.0. If the COD has a value of 0.0, there is no correlation between the observed and the forecast patterns.

The COD, when taken alone as a skill score, ignores two non-pattern related errors, the multiplicative error (MULTE) and the additive error (ADDE). The MULTE describes

an error in the amplitude of the forecast variation and can be corrected by scaling the forecast. The ADDE describes the additive bias and can be corrected by shifting the entire forecast pattern positively or negatively. Both the MULTE and the ADDE can be corrected by the MOS method outlined above. The SS then is more enlightening than the COD because it does not ignore the multiplicative and additive errors of the model, which can be corrected by post-processing techniques. A thorough discussion of the merits of the coefficients of correlation and determination as a skill score are found in Murphy (1995).

CHAPTER 4

VERIFICATION RESULTS

4.1 Coarse-mesh Findings

The coarse-mesh bar charts included in the Appendix illustrate spatial averages of each particular statistic over the domain. Values from approximately eighty sounding sites went into each spatial average for upper-air statistics and values from approximately 750 surface METAR sites went into each spatial average for surface statistics. During this warm-season study period, from 1 June to 28 August 1997, each sounding and METAR site produced approximately 47 to 60 observations for use in the verification. The forecast time $t = t_0 + 00$ h corresponds to the 0000 UTC time.

4.1.1 Temperature

Figure 1 illustrates coarse-mesh temperature verification statistics averaged over the coarse-mesh domain. Surface bias is very close to zero at the analysis time, $t = t_0 + 00$ h, growing to a positive bias (1.27 to 0.99 °C) at the early morning $t = t_0 + 12$ h and $t = t_0 + 36$ h forecast times. A negative surface bias (-1.67 °C) exists at the evening $t = t_0 + 24$ h time. In a wintertime study over the intermountain west, Manning and Davis (1997) also found that the MM5 boundary layer, using the Blackadar (1979) planetary boundary layer (PBL) scheme, was as a whole, too warm at the early morning forecast times and too cool during the early evening. In contrast a springtime verification study of MM5 over the intermountain west shows warm biases of approximately 3.0 °C in the early morning and 1.9 °C in the early evening (Kirby L. Cook, University of Utah 1998, personal communication). Nevertheless, even in the spring, the bias is high in the early morning

relative to the early evening, as in the summer and winter studies.

There are several possible explanations for the positive early morning surface temperature biases. Some of the most probable causes are: incorrect cloud cover and associated radiative effects, inaccurate surface parameterization including moisture availability, and excessive surface wind speeds. The surface temperature bias could be due to excessive vertical diffusion. Because the Blackadar PBL scheme reduces PBL depth to zero in the early morning hours, the problem would lie in the background (minimum) diffusion rather than the Blackadar scheme. Another partial explanation for the above biases could be that model surface forecast values are extracted from the lowest sigma level, which is located approximately 30 meters above ground level. Because surface observations are taken at 2 m above ground level, a perfect early morning forecast for the 30-m level could be warmer by several degrees due to diabatic effects associated with the nocturnal inversion. Throughout this paper the above altitude difference explanation will be referred to as "the altitude difference theory".

The early evening surface cool bias may be caused by all of the explanations above and also by the possibility of excessive Blackadar boundary layer turbulent transport (Zhang and Anthes 1982), which acts to destroy the surface super-adiabatic layer that should occur in nature under conditions with strong insolation over land. Additionally, the altitude difference theory may contribute to the early evening surface cool bias because a perfectly forecast early evening super-adiabatic profile would have a 30-m temperature that is several degrees cooler, in some cases, than the temperature forecast for 2 m above

ground level.

The negative temperature bias at 850 mb (-0.14°C at $t = t_0+00$ h increasing to -1.21°C with forecast time) might indicate over-prediction of the PBL depth (Shafran 1997). Both the Manning and Davis wintertime study, and the Utah springtime study find a similar negative temperature bias at 850 mb for all forecast times. There is, however, nearly zero bias at 700 mb and 500 mb for all forecast times and a slight warm bias at 250 mb for all forecast times, as found in the two prior studies. Temperature MAE values in this study are between 2.11°C and 2.52°C at the surface and decrease with height. It is possible that errors are larger at the surface because of larger local variability of surface data, differences in model and actual terrain heights and because complex boundary layer processes are difficult to parameterize.

The coefficient of determination, Figure 1 panel (e), shows that if MOS were utilized, the SS would be approximately 0.6 for the surface and 0.8 for upper-air levels. A look at the SS in panel (d) shows the model is forecasting close to that potential for 700 mb and above. Figure 1 panel (g) shows that both the surface and 850-mb levels suffer from mostly an additive error. Applying even the simplest possible MOS scheme to model temperature simulations would thus greatly improve the MM5 surface temperature forecasts.

A horizontal plot of the surface temperature bias at the early morning forecast time of $t = t_0+12$ h shows that the biases are not random, but have a distinct pattern (Figure 2). The northeastern region of the domain appears to have almost no bias, the midwest and

southern regions have a warm bias, the Rocky mountain region has a variable warm and cool bias, while the warmest bias is found along the West Coast. The warm west-coast surface bias is present for all forecast times (not shown). Some of the largest magnitudes for the surface MAE and RMSE are found within the higher elevations western regions of the domain for all forecast times (not shown). The 850-mb cool bias is also most pronounced over higher terrain areas. It is possible that this indicates inadequate terrain resolution. The SS horizontal plots (not shown) confirm that the poorest model surface-temperature forecast skill is located along the West Coast and the highest values are located in the northeastern parts of the domain.

4.1.2 Dew Point

Figure 3 shows coarse-mesh dew point verification statistics averaged across the coarse-mesh domain. Panel (a) shows an early evening surface dry bias of between -1.82 and -2.25 °C. In contrast, Manning and Davis (1997) show a morning and evening surface moist bias for the MM5 over their intermountain-west domain and an evening moist bias over the central domain. The University of Utah study finds an early evening surface moist bias of 1.8 °C and a morning surface moist bias of 1.0 °C.

One explanation for the early evening surface dry bias in the present study is that the fixed moisture availability, defined in the real-time MM5 as a function of land use, is too dry on average. This error would result in insufficient surface evaporative flux. However, while increased soil moisture might increase surface evaporation, it would also decrease the surface sensible heat flux and would thus tend to increase the magnitude of the late-afternoon cool bias found in the MM5 surface temperatures. An explanation for a small

fraction of this dry surface bias is that if the model simulation is extracted approximately 30 meters above the observed level, a perfectly modeled sounding profile would, on average, be slightly drier at that level (Fairall et al 1996).

The 250-mb dry bias of -2.54°C at the analysis time may be due to the way model moisture is handled during the dynamic initialization. Due to limited observational data, all levels above 300 mb are initialized with 10% relative humidity. As shown at the subsequent forecast times, the model quickly corrects this dry bias by transporting moisture into the upper atmosphere. The effects of initializing with a larger moisture value at upper levels should be investigated for future model runs.

The MAE of the dew point ranges from 2.04°C to 3.19°C at the surface and from 5.13°C to 7.14°C at 500 mb. Interestingly, the surface and 250-mb level MAE and RMSE values improve over time, possibly in response to the decreasing bias. At all other levels, these errors steadily increase with time. These results suggest that there may be significant errors in the model's vertical transport to the mid-troposphere.

The COD shows that the model has its greatest potential for skill at the lower levels. The surface SS suffers from mostly additive error while the 250-mb SS suffers from mainly a multiplicative error, but also a large additive error. Because the COD is much higher than the SS, MOS would greatly improve the dew point forecasts being produced by the real-time MM5.

The two-dimensional surface plots of SS, bias and RMSE (not shown) reveal that the least skillful dew-point forecasts are in the higher elevations of the Rockies. Figure 4 on

the other hand shows some of the highest 700-mb dew point SS values to be located in the same high-elevation areas. The 250-mb plots (not shown) reveal the most pronounced dry bias to be along the West Coast, close to the upwind lateral boundary.

4.1.3 Relative Humidity

Figure 5 contains coarse-mesh relative humidity (RH) verification statistics averaged across the coarse-mesh domain. The dry surface RH bias, seen in panel (a), of approximately -7% at early morning forecast times is also reported by Manning and Davis (1997) as well as the University of Utah study. Using typical values for early morning temperature and dew point, we see that a 1.3 °C rise in forecast temperature due to model bias, with no dew-point change, would translate into an RH bias of approximately 5 °C. In a similar way, the positive early morning 850-mb RH bias of about 3.5% is possibly linked to the negative temperature bias and positive dew-point bias at that level. The University of Utah study found an 850-mb positive RH bias of about 4.0% to 5.0% for all forecast times. The $t = t_0 + 00$ h dry bias at 250 mb of -2.74% helps quantify the dry initial conditions above 300 mb. Of course, the 12-h dynamic initialization may help to minimize the effect of the data paucity above 300 mb, so that most of the dry bias has already been overcome by the $t = t_0 + 000$ h. Interestingly, at the $t = t_0 + 24$ h and $t = t_0 + 36$ h forecast times the model has a moist bias of 3.99% and 2.39% respectively at the 250-mb level. RH MAE values for all levels are between 10.99% and 20.12% with the smallest error at the surface. This trend of increasing RH MAE with height is surprising since boundary layer physical processes can be difficult to simulate and the surface evaporation is based on constant values for soil moisture availability. A possible explanation could like in

errors of the vertical moisture transport of the model through convective and resolved-scale vertical velocities.

The COD shows a potential for skill of between 0.11 and 0.51. The low SS suggests that incorporating MOS into the model would significantly improve the real-time MMS RH forecasts. Most of the SS deficiencies are due to additive errors at lower levels, but multiplicative errors are dominant at 250 mb.

4.1.4 Wind Direction

Figure 6 displays coarse-mesh wind direction verification statistics averaged over the coarse-mesh domain. A positive bias means that the forecast wind direction was veered (clockwise) with respect to the observed wind direction and a negative bias means that the forecast wind direction was backed (counter-clockwise) with respect to the observed direction.

The early-morning surface veering or positive bias of 12.67 to 14.95 degrees might be explained by the excessive background vertical diffusion possibility mentioned in section 4.1.1. The vertical diffusion would tend to pull, faster and veered winds down toward the surface from aloft. The early morning surface veering also can be explained in part by the altitude difference theory because the model surface layer wind values come from 30 meters above ground level where the winds would be slightly veered. At and above the 850-mb level the small directional biases are between 2.0 and 4.0 degrees. The MAE values are between 42.88 and 51.38 degrees for the surface and improve with height to between 10.24 and 23.55 degrees at 250 mb. The large direction errors at the surface in

the early morning are associated with the local variability that is characteristic of the slower wind speeds commonly observed at that time.

Two-dimensional surface wind direction bias plots (not shown) indicate that all portions of the domain contribute evenly to the early morning veering of surface wind. There are a large number of model surface sites in the western half of the domain, where terrain complexity is greatest, with large wind-direction bias magnitudes. Surface RMSE values also are higher for the western half of the domain due to the more complex terrain.

4.1.5 Wind Speed

Figure 7 shows coarse-mesh wind speed verification statistics averaged over the coarse-mesh domain. The positive 2.2 m s^{-1} surface wind speed bias for early morning forecast times might be explained by the excessive background vertical diffusion theory because faster wind speeds from higher levels would be transported down to the surface. The altitude difference theory may explain some of the speed bias because wind shear in the lowest 100 m can be large at night. At the $t = t_0 + 24 \text{ h}$ early evening forecast time there is also a positive, yet weaker, surface speed bias of 1.01 m s^{-1} . Similar results appear in Manning and Davis (1998) and The University of Utah study as well as a study by Oncley and Dudhia (1995). The Cox et al. (1998) study shows a positive surface wind speed bias for all four of the models in the comparison. Wind speeds in this study are slightly weak for levels above the surface.

The COD values show poor potential for skill at the surface with greater potential at higher levels. The utilization of MOS would correct the significant deficiencies due to

both additive and multiplicative error in the surface wind skill score, but would be of little help at the higher levels where wind speeds are simulated well.

4.1.6 Geopotential Height

Figure 8 displays coarse-mesh geopotential height verification statistics averaged over the coarse-mesh domain. The height field simulations are second only to sea level pressure in skill. The 850-mb heights have a positive bias of only 3.41 to 9.93 meters; similar results for 850 mb were found in the University of Utah study. Looking back at Figure 1 we see that the temperature of the model layer from the surface to 850 mb should be colder than observed, but we see that there is a positive 850-mb height bias. This contradiction may be due to the interpolation of the height values when converting from sigma to pressure coordinates. The 250-mb heights have a negative bias of -2.31 to -19.12 meters and intermediate levels have a near-zero bias. Similarly, the University of Utah study shows a 250-mb bias of between 0.0 and -8.0 meters with reduced bias at mid levels. For all levels the bias is a very small percentage of the typical value for geopotential height variations in time. MAE values for all levels are between 7.77 and 23.77 meters, which is also a small fraction of the typical range of variation in time.

The COD illustrates a potential for very high SS. The multiplicative and additive errors are very small, resulting in SS values that are very close to the high potential seen in the COD values. Errors are so small that utilizing MOS would not significantly improve height forecasts.

4.1.7 Sea Level Pressure

Figure 9 shows coarse-mesh sea-level pressure verification statistics averaged over the coarse-mesh domain. Biases are very small, ranging from -0.41 to 0.64 mb. MAE values range from 1.23 to 1.63 mb. The spatially averaged SS values are high overall and would not be improved significantly by MOS. These statistics indicate excellent quality in the sea-level pressures on the 36-km domain.

Clear regional biases for sea-level pressure can be found in the horizontal analyses. For example, Figure 10 shows the horizontal field of sea-level pressure biases at $t = t_0 + 12$ h. The northeast portion of the domain has the smallest positive bias, the plains and most of Canada have a small negative bias and the western region has a positive sea level pressure bias that becomes large in some areas. These regional biases may suggest a systematic problem in the reduction of the model's surface pressure forecasts to sea level.

4.2 Fine-mesh Findings

The fine-mesh bar charts included in the Appendix illustrate spatial averages for each particular statistic over the domain. Values from nine sounding sites went into each spatial average for upper-air statistics and values from approximately 100 surface METAR sites went into each spatial average for surface statistics. During this study period each sounding and METAR site produced approximately 47 to 60 observations for use in the verification. The forecast time $t = t_0 + 00$ h corresponds to the 0000 UTC time.

4.2.1 Temperature

Figure 11 illustrates temperature verification statistics averaged over the fine-mesh domain. Similar to the coarse mesh statistics that are calculated over only the fine-mesh

domain, the fine mesh has an early morning surface warm bias of 1.38 °C. Both the fine-mesh statistics and the coarse-mesh statistics, which are calculated over the fine-mesh domain, show a slight cooling trend at 850 mb. Because the fine-mesh model was initialized with a warm bias at 850 mb of 0.29 °C, the fine-mesh statistics do not show a cool bias like the coarse-mesh statistics, that are calculated over the fine-mesh domain, of between -0.15 °C and -0.5 °C. MAE values are between 1.84 °C and 1.97 °C at the surface and fall to about 0.9 °C for all other levels.

Except for the surface, the COD values (not shown) are not much larger than the SS values in Figure 11, indicating that MOS would be of only slight help for surface temperatures. The additive error is the greatest contributor to the reduction of the SS values at the surface.

For the domain-wide early morning surface warm bias, the two-dimensional spatial plot (Figure 12) indicates variable warm and cool bias over the higher terrain and an almost zero bias for many of the model locations along the coast.

4.2.2 Dew Point

Figure 13 shows dew point verification statistics averaged over the fine-mesh domain. An evening surface dry bias exists on the fine mesh, which is similar to, yet weaker than that seen for the coarse-mesh statistics averaged over the fine-mesh domain. The 700-mb level is dry by -0.98 °C to -2.08 °C, while the 500-mb level is too moist by between 1.06 °C and 1.64 °C. The 250-mb level was initialized on the fine mesh with very close to zero bias, but quickly jumps to a moist bias of between 1.9 °C and 2.14 °C. MAE and RMSE

values are greatest for the upper levels.

The COD values are much smaller for the upper levels than for lower levels (not shown). The levels with low COD values indicate a poor match of the patterns between the simulated and observed dew-point fields. SS values are low for upper levels and plunge for the 250-mb level. Most of the problem at 250 mb is in the multiplicative error and both the surface and the 250-mb level suffer from an additive error of 0.25 (not shown).

Examining the poor skill of the upper-level model dew point on the two-dimensional spatial plots, it is found that the largest errors occur near the coast. The same pattern holds true for all forecast times and levels. Figure 14 is an example of the coastal 250-mb moist bias, shown for $t = t_0 + 24$ h. This result may imply a tendency for excessive deep convection in the coastal region.

4.2.3 Relative Humidity

Figure 15 displays relative humidity verification statistics averaged over the fine-mesh domain. For all levels, RH bias is between -9.16% and 8.91% with the driest level at 700 mb and most moist at 250 mb. MAE and RMSE values increase gradually with height for levels above the surface.

The COD values are between 0.28 and 0.51 for all levels and at all times except at the 250-mb level where they are between 0.10 and 0.16 (not shown). The surface early morning negative value in the SS is primarily due to an additive error, and the SS for 250 mb is very low and suffers mostly from a multiplicative error (not shown).

Two-dimensional relative humidity bias figures (not shown) indicate that the early morning surface dry bias is wide spread over land with a strip of almost zero RH bias along the coast. As for the dew point the coastal areas gave the largest contribution to the moist bias error at 250 mb for all forecast times (not shown).

4.2.4 Wind Direction

Figure 16 illustrates wind direction verification statistics averaged over the fine-mesh domain. At the early morning forecast time $t = t_0 + 12$ h, surface winds are veered or have a positive bias with respect to the observed direction. This veering may be explained by the excessive background vertical diffusion theory since vertical diffusion within the lower atmosphere would serve to transport faster and more veered winds down to the surface. Interestingly, the early evening model wind directions are backed with respect to the observations for the 850, 700 and 500-mb levels by 5 to 7 degrees. This is the most consistent deviation from geostrophy seen in the statistics. Domain-average MAE values for surface directions are between 39 and 41 degrees at the surface. Upper level MAE values are between 11 and 26 degrees with error values improving gradually with height. No skill scores were calculated for wind direction.

Two-dimensional fields of surface wind-direction bias at both the early morning and early evening forecast times indicate consistent negative bias or backing along the east coast (not shown). One possible explanation for this is excessive surface friction, that would cause the surface winds to back (shift counter-clockwise) from the prevailing warm-season southerly wind direction along the coast (E.S.S.A. 1968).

Additionally, the surface bias plots (not shown) at both early-morning and late-evening forecast times reveal that in West Virginia, on the western side of the Appalachians, the winds are veered, while in Virginia, on the eastern side, they are backed, relative to the observations. This veered-backed couplet could indicate excessive convergence over the mountain ridges during the day, possibly related to the local thermal pattern.

4.2.5 Wind Speed

Figure 17 displays wind-speed verification statistics averaged over the fine-mesh domain. There is a consistent positive wind speed bias at the surface for all times of about 2.0 m s^{-1} which could be explained in part by the altitude difference theory because model winds are usually faster 30 meters above the surface. Wind speed MAE values are approximately 2.0 m s^{-1} at mid levels and approximately 3.5 m s^{-1} at 250 mb which are relatively small compared to average wind speeds at those levels. In contrast, the surface MAE values are between 2.3 m s^{-1} and 2.8 m s^{-1} , which are about equal to the average surface wind speed variation in time and in space. Consequently model SS values are very low at the surface, but improve markedly for upper levels. The lack of skill at the surface is due to two parts additive error and one part multiplicative error (not shown). Two-dimensional fields of surface wind-speed bias indicate that the overall positive bias, illustrated in panel (a) of Figure 17, is almost independent of location (not shown).

4.2.6 Geopotential Height

Figure 18 shows geopotential-height verification statistics averaged over the fine-mesh domain. Panel (a) shows no significant biases at any level and for any time. MAE values are between 6.05 and 18.69 m and SS values are very high for all times and all levels.

Multiplicative and additive errors (not shown) are both below 0.02 for all levels and all times indicating that MOS would not significantly improve these geopotential height simulations. Two-dimensional fields (not shown) indicate consistently higher RMS error values and lower SS values for geopotential height in the eastern portion of the domain.

4.2.7 Sea Level Pressure

Figure 19 displays sea-level pressure verification statistics averaged over the fine-mesh domain. The bias has its largest negative at the analysis time (-0.9mb), but improves for the two forecast times (-0.2mb and 0.3mb). MAE and RMS Error values are only about 1 mb for all times and SS values are consistently close to one. MOS would not significantly improve these skillful fine-mesh SLP forecasts. Two-dimensional fields of the SLP bias showed the highest magnitudes are located over the Appalachians. Fortunately, the range of bias errors on the fine-mesh grid is only from -1.0mb to 0.7mb. There is a noticeable positive bias of about 0.5mb along the coast of the fine-mesh domain for the $t = t_0 + 12$ h and the $t = t_0 + 24$ h forecast times (not shown).

4.3 Coarse and Fine-mesh Model Vertical Structure Verification

In an attempt to quantify the quality of the vertical structure of the forecasts on coarse and fine-mesh models, a domain-wide average of the bias and RMSE of the total totals convective index is calculated. The total totals index (TTI) is used as an aid in forecasting severe convective weather and is defined as

$$TotalTotalsIndex = T_{850mb} + TD_{850mb} - 2 \cdot T_{500mb} \quad (8)$$

where T is temperature. The higher the value of the TTI, the larger the probability of

severe weather. TTI values below 50 indicate weak probability of severe convective weather, values between 50 and 55 indicate moderate probability, and values above 55 mean there is a good probability of severe convective weather (Doswell 1986).

Values of the bias of TTI, shown in Figure 20 panel (a), are about -1.0 for the fine mesh and -1.5 for the coarse mesh for all times displayed. Values for the RMSE of total totals index, shown in Figure 20 panel (b), are about 4.0 units for the fine-mesh domain and 5.0 units for the coarse-mesh domain for all times displayed. The fine-mesh advantage of lower RMS Error and smaller bias could be significant to a forecaster facing a decision on issuance of a severe weather warning or advisory.

4.4 Coarse Mesh Versus Fine Mesh

For the purpose of directly comparing coarse-mesh statistical values with fine-mesh statistical values, we calculate an average of the coarse-mesh verification statistics over the smaller geographical area of the fine mesh. The comparison results are displayed in Figure 21. Because of uncertainty in the degrees of freedom for data sets containing back to back forecasts, we make no attempt to determine if the findings of the comparison are statistically significant. Most of the differences between the coarse mesh and fine-mesh statistical values, however, are very small.

There are 248 black boxes and 97 white boxes in Figure 21, indicating that the coarse-mesh produced forecasts over the Northeast that were better than or equal to the fine mesh for more than two out of three of the statistical measures. What cannot be seen by these discrete “win or lose” boxes in Figure 20 is the fine-mesh skill improvement over

time. For some of the variables, the fine-mesh bias is larger than that of the coarse-mesh at the analysis time ($t = t_0 + 00$ h) but quickly improves, and almost surpasses the coarse mesh, for subsequent forecast times. Similarly, initial RMSE, MAE, and SS values for the fine mesh are often greater than the coarse-mesh values, but generally improve with each successive forecast time.

Because the fine mesh frequently has higher errors and lower skill than the coarse mesh at t_0 , the analysis time, it is inferred that the fine mesh needs an improved initialization procedure. Currently, the fine mesh initial conditions are interpolated from the coarse mesh dynamically initialized values at the $t = t_0 - 3$ hour startup time. No data are assimilated into the fine mesh during this 3-h period, except via the coarse-mesh lateral boundary conditions. These results suggest that the 12-h coarse mesh dynamics initialization, based on nudging, is having a lasting positive effect on the coarse-mesh forecast skill. A better initialization procedure needs to be developed that allows the fine mesh to spin up model moisture and divergence while ingesting new observations right up to the $t = t_0 + 00$ h analysis time. With an improved initialization, model bias, error and skill should be about the same for both the fine and coarse meshes at analysis time. If the fine-mesh error and skill scores improve over time as they do now, the fine mesh forecast values should out-perform those on the coarse mesh for several variables, especially the moisture fields.

CHAPTER 5

CONCLUSIONS

5.1 Summary of MM5 Strengths and Weaknesses

The coefficient of determination values show good overall pattern matching for all variables at all levels, except for the 250-mb moisture and the surface wind speed. This result shows that the model has high potential for skill, especially if multiplicative and additive errors are eliminated through post-processing, such as MOS. Evening surface temperature, morning surface dew point, and sea-level pressure, as well as upper-level temperature, wind speed, wind direction, and geopotential height, all score well with the current model configuration. On average the portion of the coarse-mesh model domain east of the Rockies shows the smallest error and largest skill score values.

In contrast, early morning surface temperature, RH, wind direction, and wind speed bias values, as well as evening surface temperature and dew point bias values, show notably reduced skill compared to those listed above. There are several possible explanations for the positive early morning surface biases. Some of the most probable causes are: incorrect cloud cover and associated radiative effects, inaccurate surface parameterization including moisture availability, and excessive surface wind speeds. Because the Blackadar PBL scheme reduces PBL depth to zero in the nocturnal period, the early-morning surface temperature bias could be due to excessive background vertical diffusion.

Another partial explanation for the biases in surface wind speed and direction

could be that the model's surface forecast values are extracted from the lowest sigma level, which is located approximately 30 meters above ground level. Because surface wind observations are taken at about 10 meters above ground level, this altitude difference might cause a portion of the positive bias in MM5 wind direction and speed.

The early evening surface cool bias may be related to all of the factors above and the possibility of excessive turbulent transport by the Blackadar boundary layer which would serve to destroy the surface super adiabatic layer that should occur in nature during strongly heated conditions.

A coarse-mesh 850-mb cool bias is present at all forecast times possibly indicating an over-prediction of boundary layer depth. There is an 850-mb warm initialization, followed by a cooling trend in the fine-mesh statistics, possibly indicating that the same mechanism that is producing the 850-mb cool bias shown in the coarse mesh statistics could be present in the fine-mesh model. Both the coarse and fine mesh demonstrate reduced skill for temperature predictions at the 250-mb level.

For most variables and most times, two-dimensional fields of the statistical results reveal that the West Coast and mountainous regions typically have the largest errors and smallest skill scores. Reduced skill on the West Coast may be due to its proximity to the western model boundary. Reduced skill in the mountainous regions could be due to smoothed representation of the terrain or limited grid resolution. Because of the regional variability in skill observed in all two-dimensional fields, application of MOS would be able to improve the predictions for all variables in at least a portion of the domain.

Based in part on the findings in this study, several changes are being made to the next version of the Penn State real-time MM5 forecast system. To improve the prediction of surface super-adiabatic layers, a turbulent kinetic energy (TKE) based PBL scheme will replace the Blackadar PBL. To improve West-Coast forecasts, the 36-km grid will be expanded toward the west and a new 108-km grid will be added that extends far into the Pacific Ocean. Additionally, the latest version will include dynamic initialization on the 12-km model grid using analysis-nudging FDDA to help improve the fine-mesh initialization.

In general the coarse mesh exhibits better skill than the fine mesh when we calculate coarse-mesh statistics over the geographic boundaries of the fine-mesh domain, especially at the initial time. A closer look at the scores, however, shows that the fine-mesh skill rapidly improves for many variables as the forecasts proceed. An improved fine-mesh initialization procedure could allow the fine-mesh analysis to have equal skill as the coarse-mesh analysis, so that the subsequent fine-mesh forecast skill might surpass the coarse-mesh skill.

We also verified the total totals index as a first step in assessing the quality of model vertical-structure forecasts, which are important for predicting deep convection. Results show a consistent negative bias of total totals index for both the coarse and fine meshes. For all forecast times, the fine mesh has a bias of approximately -1.0 units while the coarse mesh has a larger error of -1.5 units. Fine-mesh RMSE values are approximately 4.0 units and coarse-mesh RMSE values are larger at approximately 5.0

units for all forecast times. These biases and RMSE values are small enough that a forecaster can use MM5 total totals simulations as an aid for severe convective weather prediction. The finer resolution of the fine-mesh domain might account for the better skill at predicting the total totals index.

5.2 Future Work

To give a more complete and accurate verification of the model, a more complete set of convective indices describing the vertical structure of the mass field (stability) needs to be evaluated in addition to total totals index. The set should include: convective available potential energy (CAPE), convective inhibition, downdraft CAPE, BEST CAPE (largest realizable), lifted condensation level, planetary boundary layer depth, moist conditional symmetric instability and storm relative helicity.

Future verification of precipitation forecasts will test the skill of the model's convective parameterizations, explicit moist physics, vertical thermodynamic structure, and precipitation event timing, among other important characteristics. The observations that will be used for the precipitation verification are the StageIV, gauge-corrected, 4km-resolution precipitation data being produced by the Environmental Modeling Center at the National Weather Service. Threat and bias scores will be calculated and new techniques for objectively verifying mesoscale precipitation will be developed that can take spatial and temporal errors into account.

REFERENCES

- Anthes, R.A., 1983: Regional Models of the Atmosphere in Middle Latitudes. *Monthly Weather Review*, 111, 1306-1335.
- Blackadar, A. K., 1979: High-resolution models of the planetary boundary layer. *Advances in Environmental Science and Engineering, No. 1*. Pfafflin and Ziegler, Eds., Gordon and Breach Sci. Publ., New York, 50-85.
- Brier, G. W., and R. A. Allen, 1951: Verification of Weather Forecasts. *Compendium of Meteorology*, T.F. Malone, Ed., American Meteorological Society, 841-848.
- Cox, R., B. L. Bauer, T. Smith, 1998: A Mesoscale Model Intercomparison. *Bulletin of the American Meteorological Society*, 79, 265-283.
- Doswell, C. A., 1986: Short Range Forecasting. *Mesoscale Meteorology and Forecasting*, P. S. Ray, Ed., American Meteorological Society, 793 pp.
- Dudhia, J., 1993: A nonhydrostatic version of the Penn State / NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Monthly Weather Review*, 121, 1493-1513.
- Environmental Sciences Services Administration, 1968: Climatic Atlas of the United States. U. S. Department Of Commerce, U. S. Government Printing Office, Washington DC, 80pp.
- Fairall, L. W., E. F. Bradley, D. P. Rogers, J. B. Edson, and G. S. Young, 1996: Bulk Parameterization of Air Sea Fluxes for TOGA COARE. *Journal of Geophysical*

Research, 101, 3734-3764.

Glahn, H. R. and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *Journal of Applied Meteorology*, 11, 1203-1211.

Grell, G. A., J. Dudhia, D. R. Stauffer, 1994: A description of the fifth-generation Penn State/NCAR mesoscale model (MM5). *NCAR Technical Note-398+STR*, National Center for Atmospheric Research (NCAR), Boulder, 138pp.

Kain, J. S. and J. M. Fritsch, 1993: Convective parameterization in mesoscale models: The Kain-Fritsch scheme. *The representation of cumulus convection in numerical models*. Meteorological Monographs, No. 46, American Meteorological Society. K. A. Emanuel, and D. J. Raymond, EDS. 165-170.

Manning, K. W., and C. A. Davis, 1997: Verification and Sensitivity Experiments for the WISP94 MM5 Forecasts. *Weather And Forecasting*, 12, 719-735.

Murphy, A. H., 1995: Coefficients of Correlation and Determination as Measures of Performance in Forecast Verification. *Weather And Forecasting*, 10, 681-688.

_____, A. H., 1988: Skill Scores Based on the Mean Square Error and Their Relationships to the Correlation Coefficient. *Monthly Weather Review*, 116 2417-2424.

_____, A. H., and H. Daan, 1985: Forecast Evaluation. *Probability, Statistics, and Decision Making in the Atmospheric Sciences*, A. H. Murphy and R. W. Katz, Eds., Westview Press, 379-437.

- _____, A. H., and R. L. Winkler, 1987: A General Framework for Forecast Verification. *Monthly Weather Review*, 115, 1330-1338.
- Oncley, S. P., and J. Dudhia, 1995: Evaluation of Surface Fluxes from MM5 Using Observations. *Monthly Weather Review*, 123, 3344-3357.
- Shafran, P. C., 1997: Summertime Circulation Patterns in the Lake Michigan Area. *MS Thesis, The Pennsylvania State University*, 206pp.
- Stauffer, D. R. and N. L. Seaman, 1990: Use of Four-Dimensional Data Assimilation in a Limited-Area Mesoscale Model. Part 1: Experiments with Synoptic-Scale Data. *Monthly Weather Review*, 118, 1250-1277.
- White, B. G., J. Paegle, W. J. Steenburgh, J. D. Horel, R. T. Swanson, D. J. Onton, and J. Miles, 1998: Short-Term Forecast Validation of Six Models for Winter 1996. *Weather and Forecasting*, in press.
- Zhang, D. L. and R.A. Anthes, 1982: A high-resolution model of the planetary boundary layer-sensitivity tests and comparison with SESAME-79 data. *Journal of Applied Meteorology*, 114, 1330-1339.

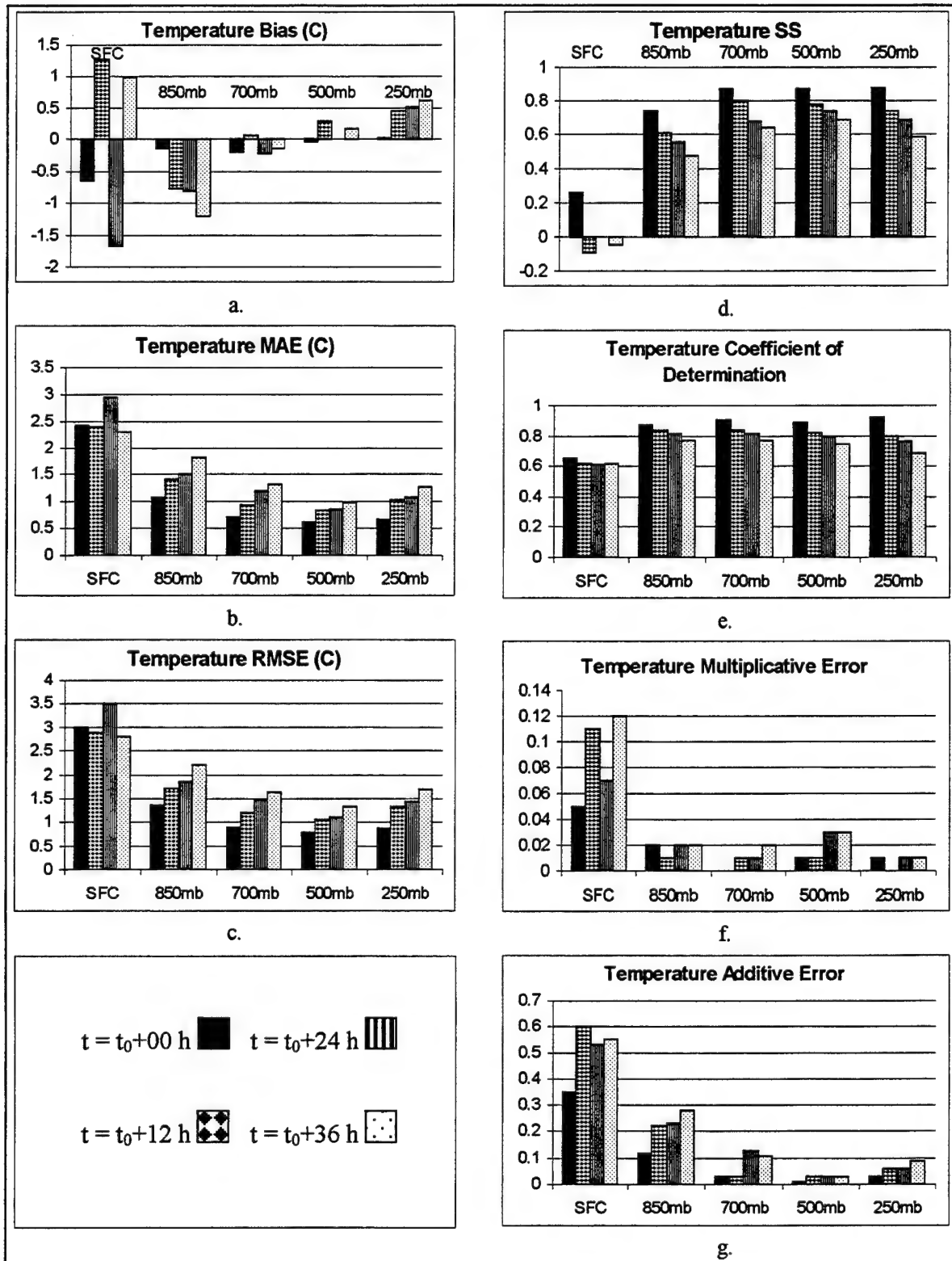


Figure 1. Coarse-mesh temperature statistics for 1 June to 28 August 1997.

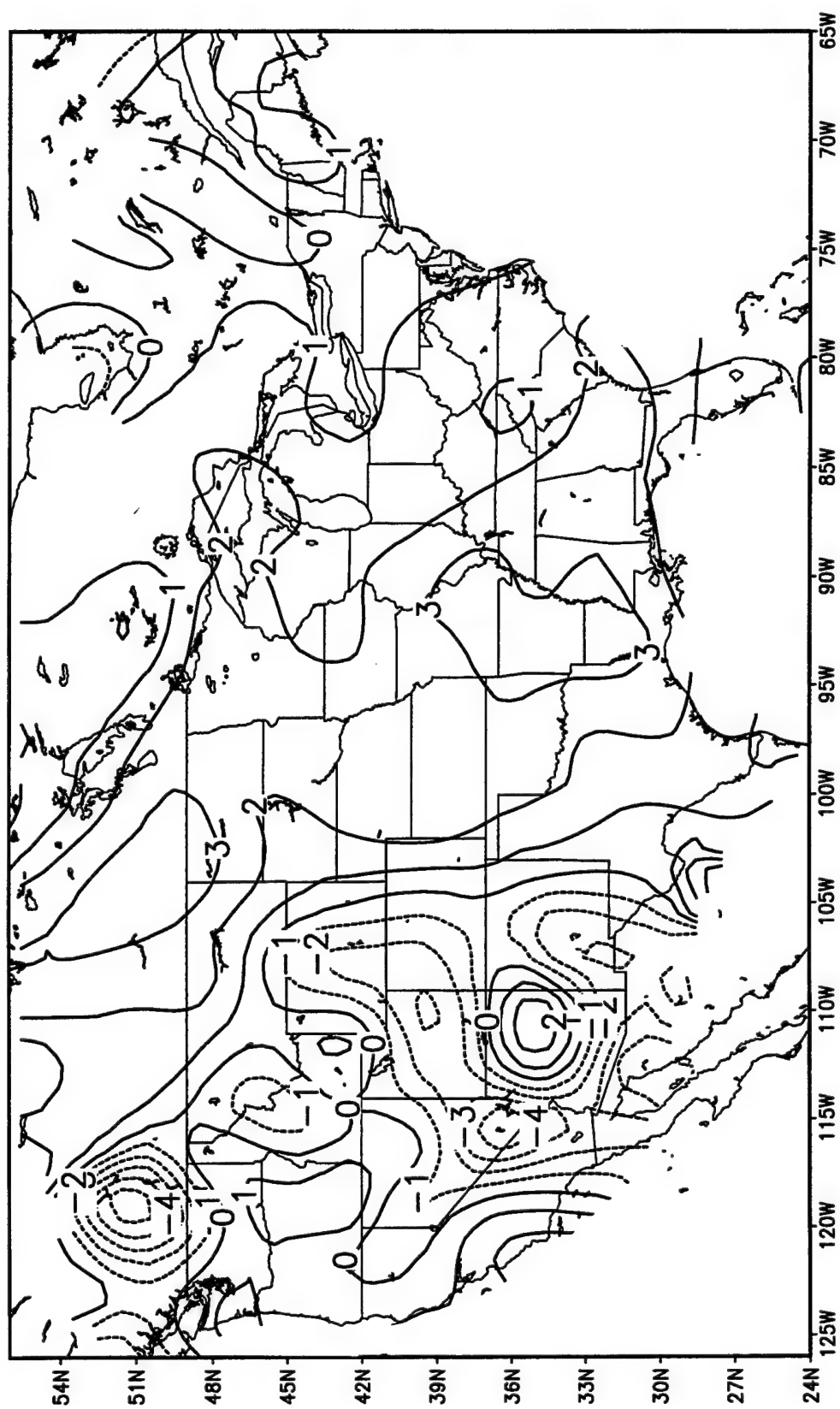


Figure 2. Surface temperature bias, forecast time = $t_0 + 12$ h (contour interval of 1 °C). 3

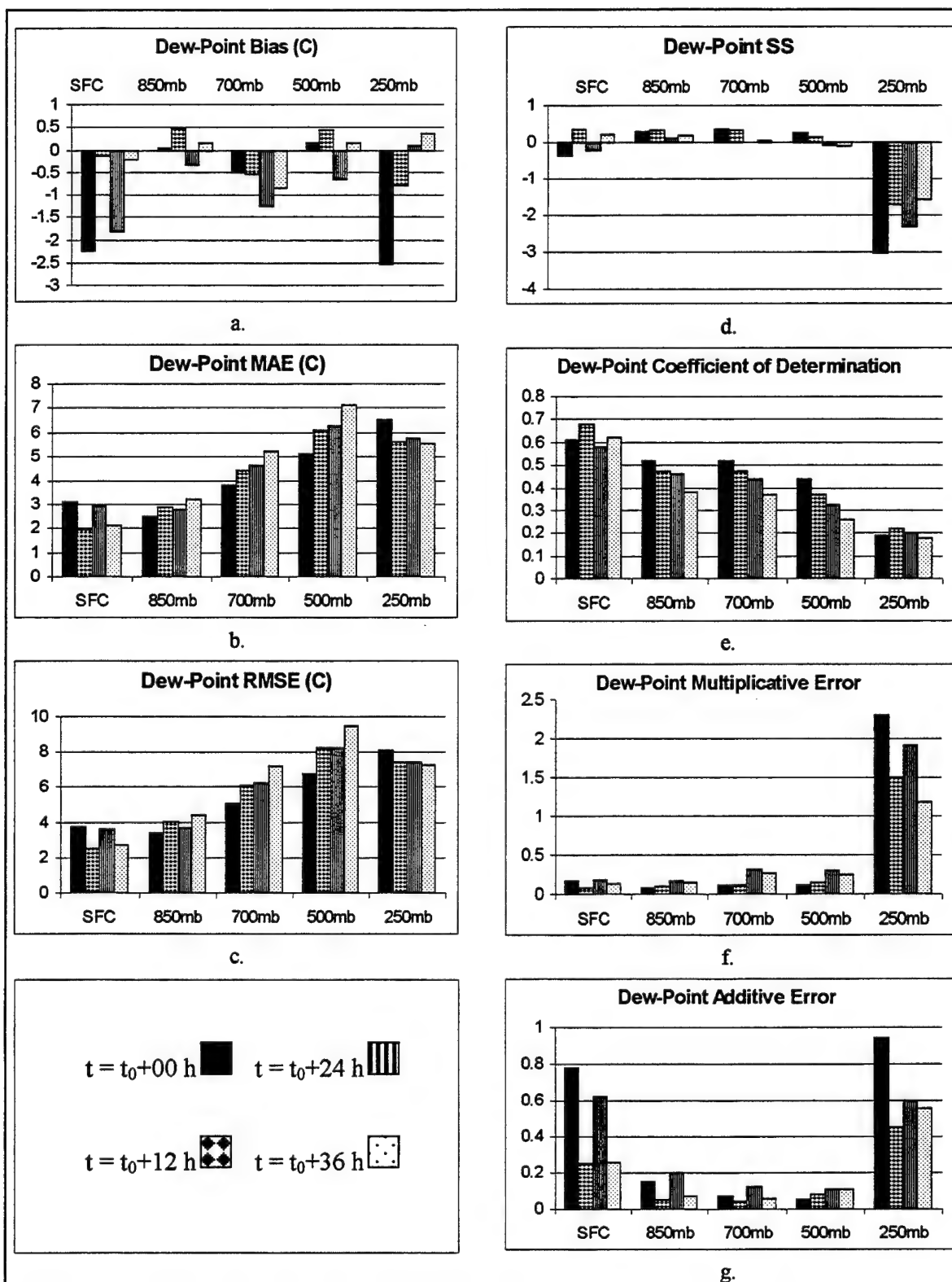


Figure 3. Coarse-mesh dew-point statistics for 1 June to 28 August 1997.

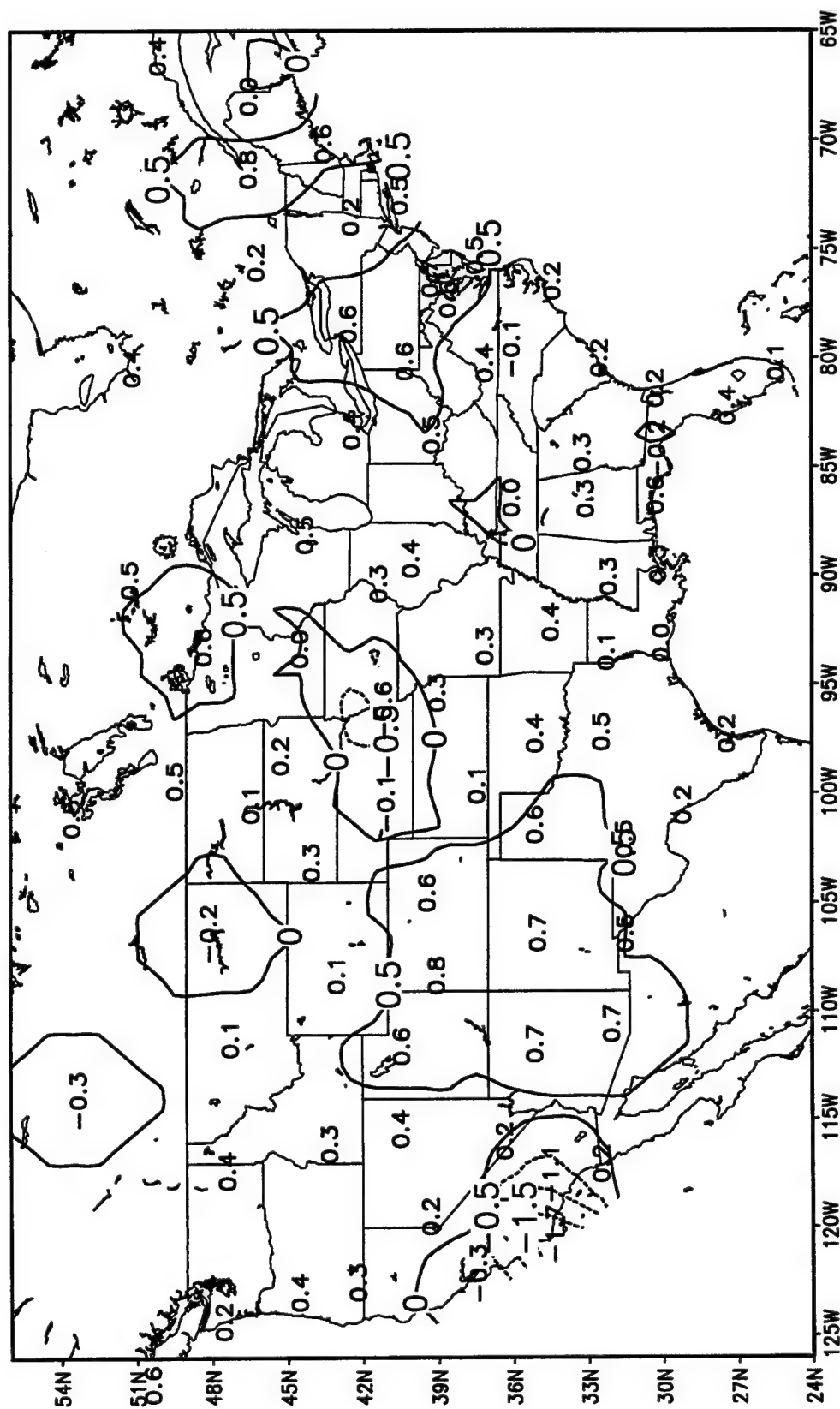


Figure 4. Dew-point skill-score at 700 mb, forecast time = $t_0 + 12$ h
(contour interval of 0.5).

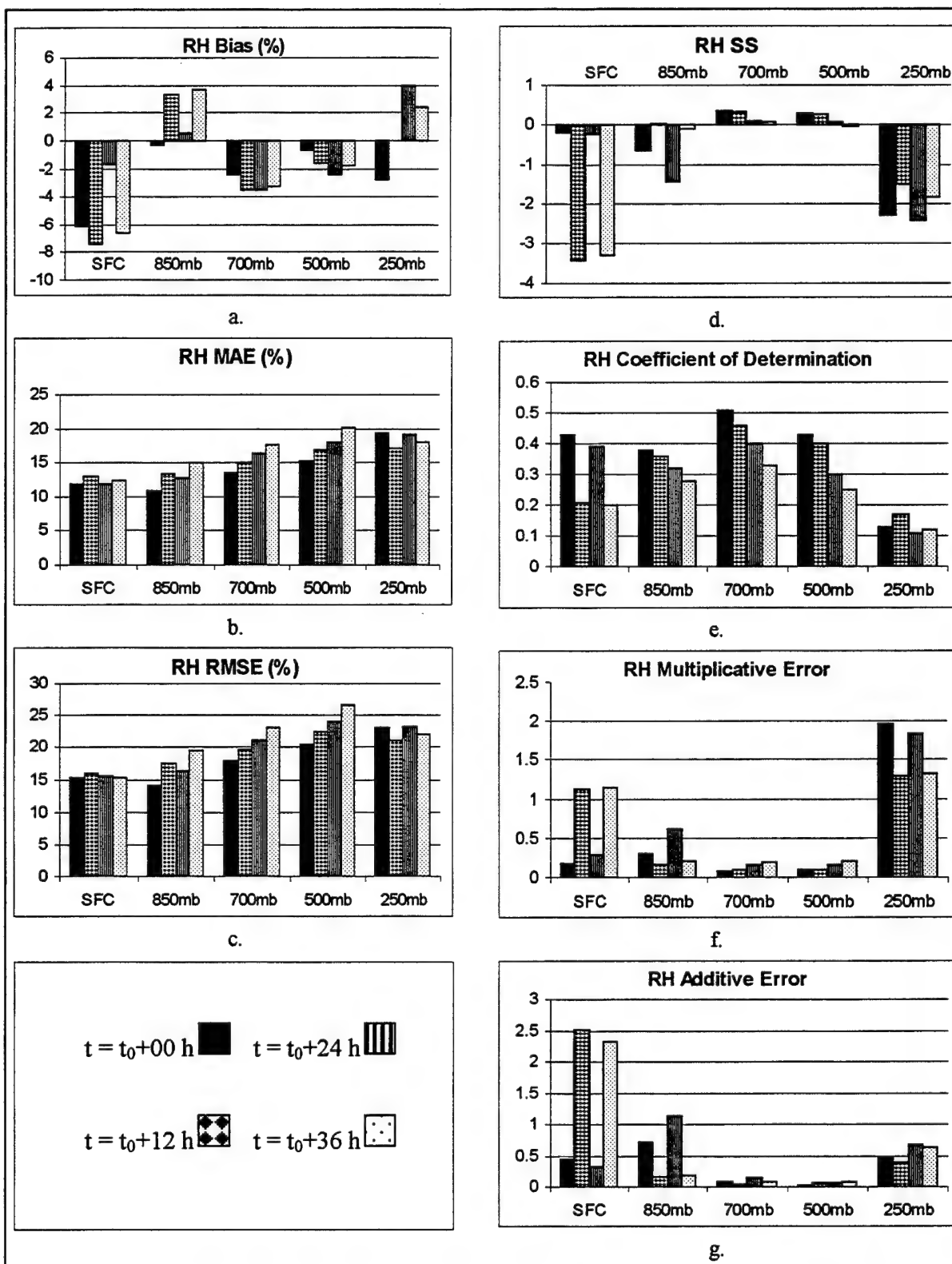


Figure 5. Coarse-mesh relative-humidity statistics for 1 June to 28 August 1997.

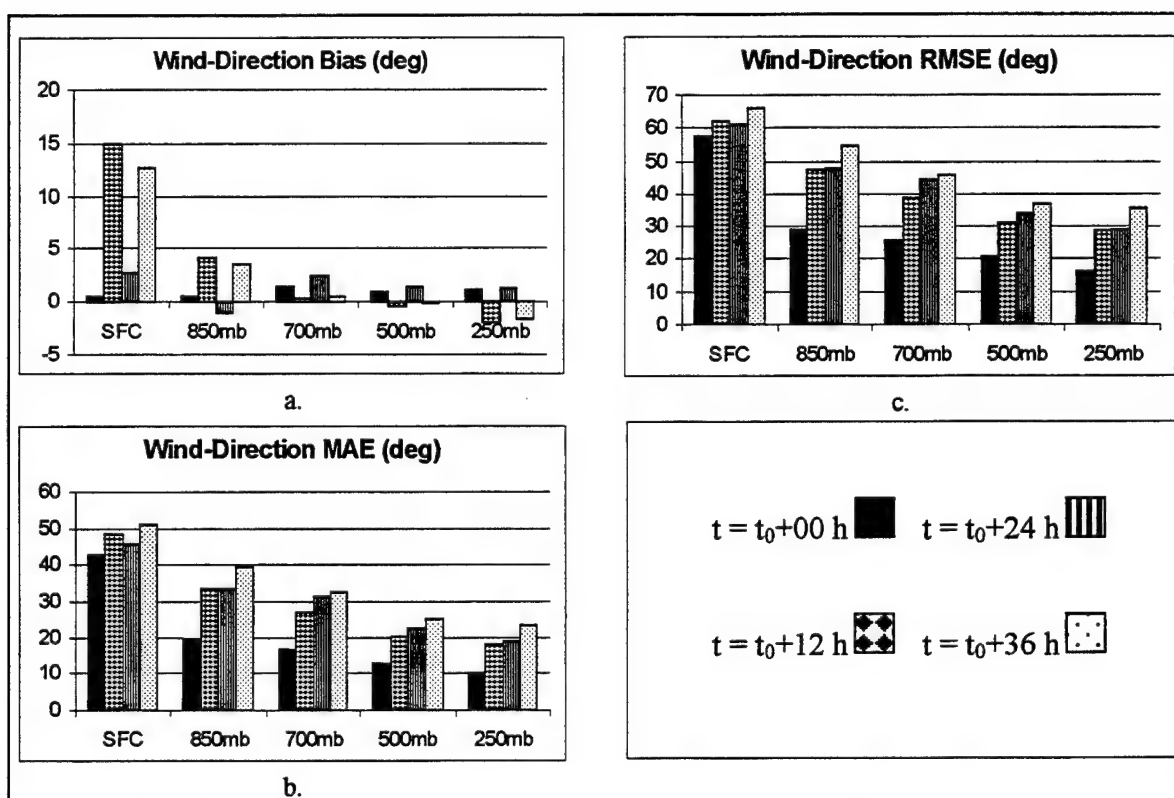


Figure 6. Coarse-mesh wind-direction statistics for 1 June to 28 August 1997.

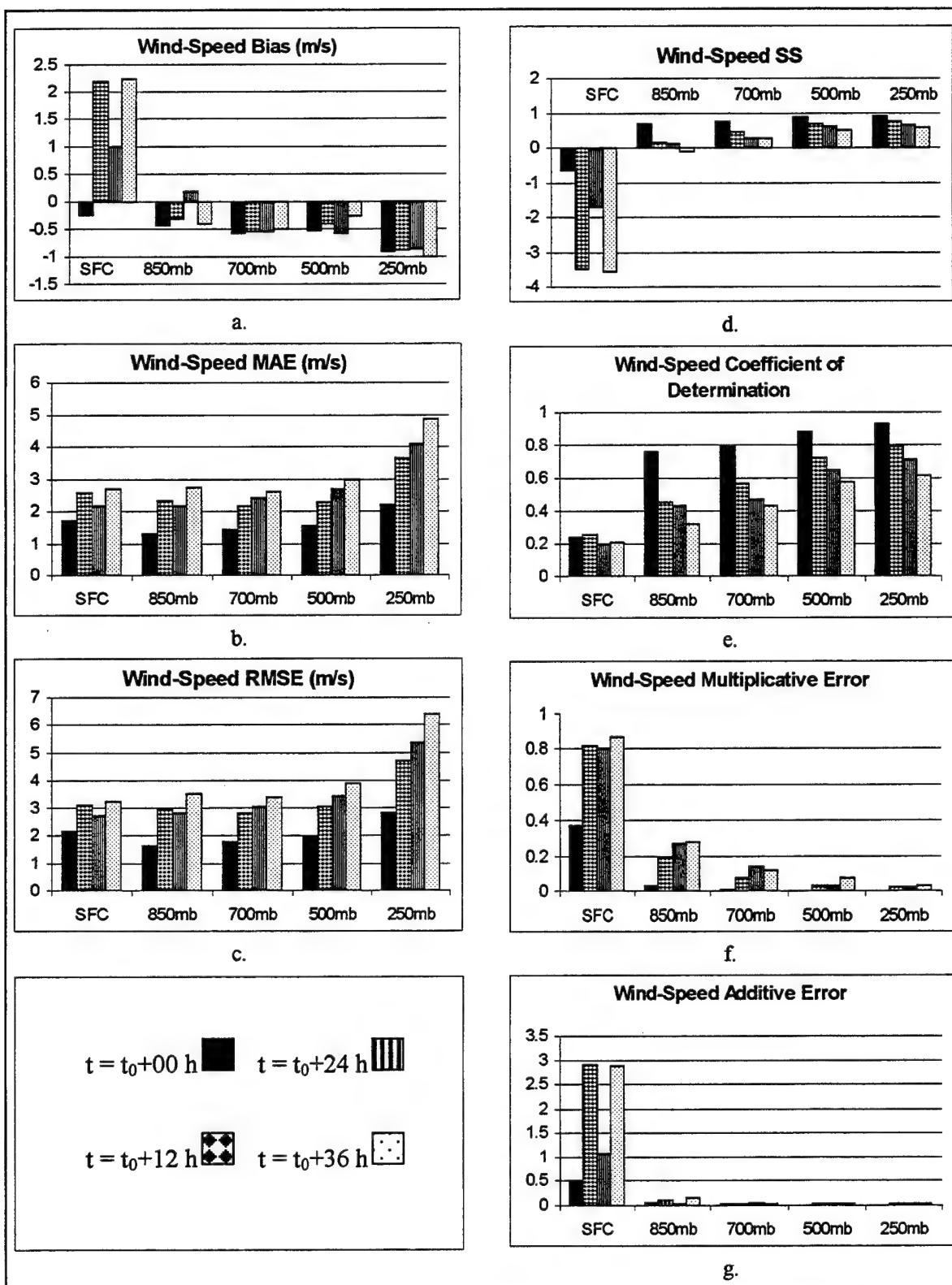


Figure 7. Coarse-mesh wind-speed statistics for 1 June to 28 August 1997.

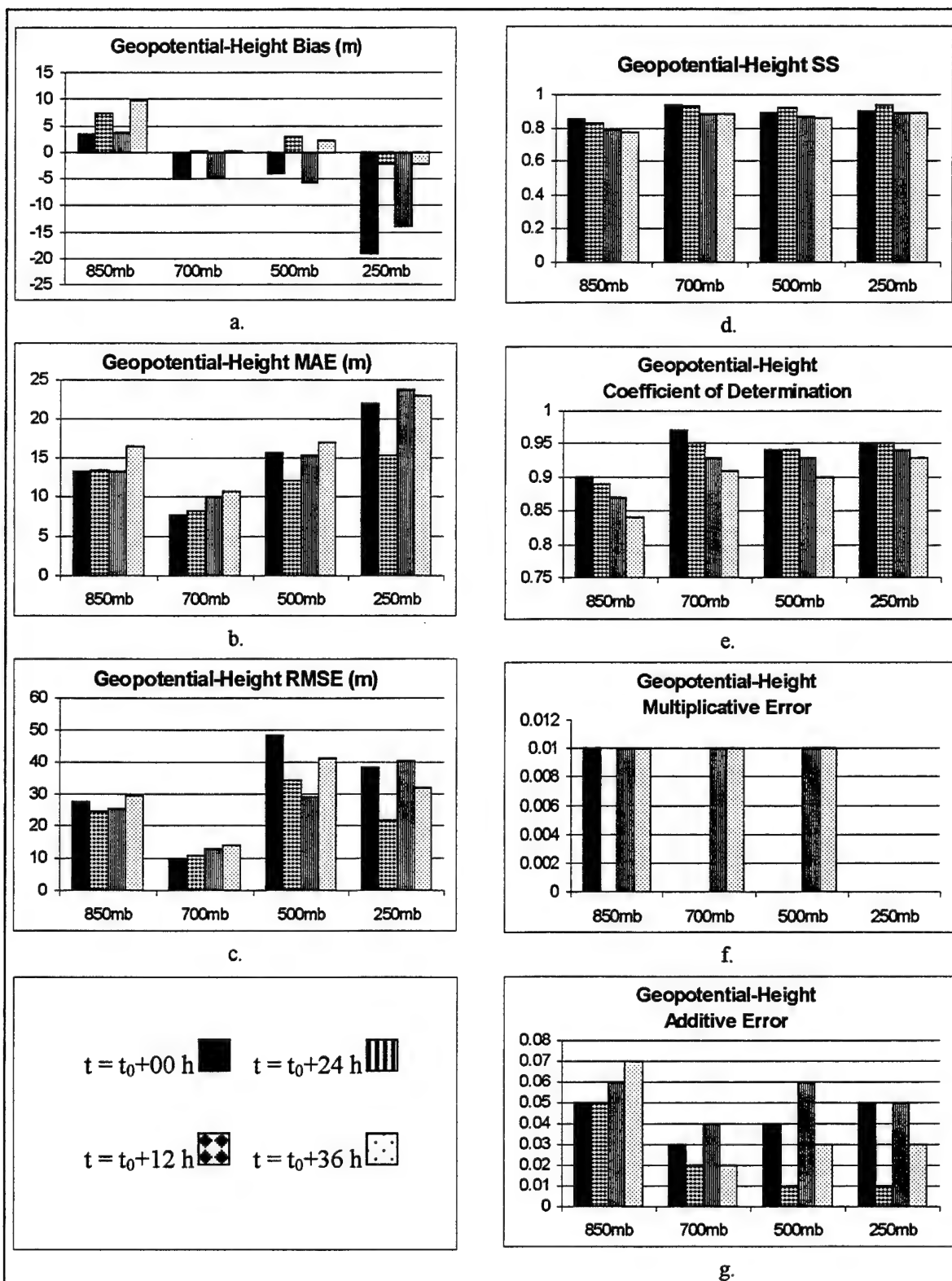


Figure 8. Coarse-mesh geopotential-height statistics for 1 June to 28 August 1997.

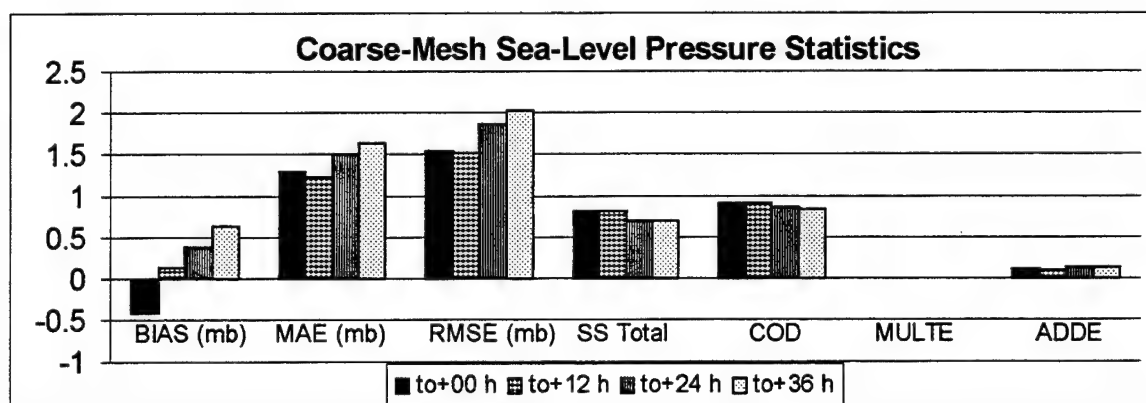


Figure 9. Coarse-mesh sea-level pressure statistics for 1 June to 28 August 1997.

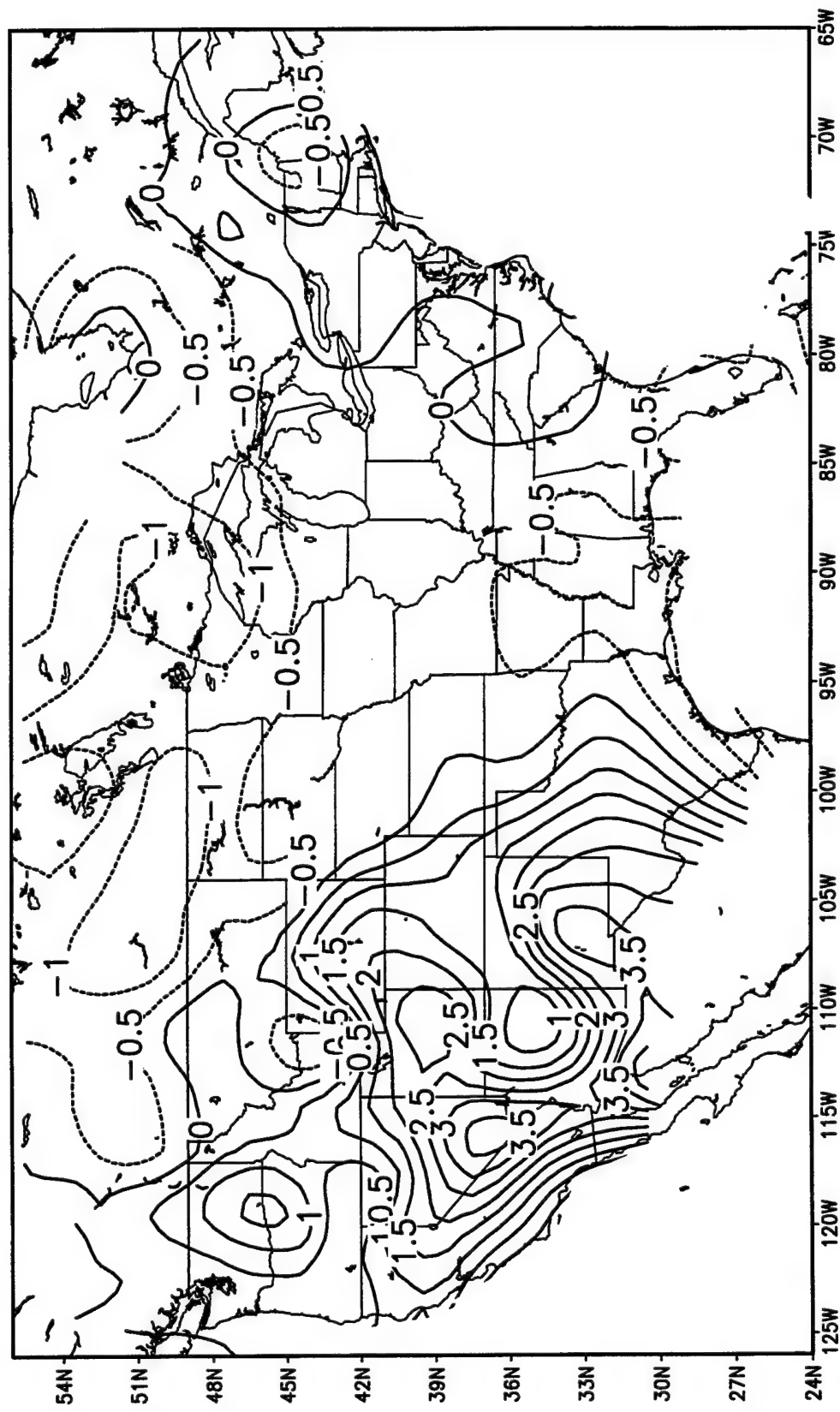


Figure 10. Sea-level pressure bias, forecast time = $t_0 + 12$ h (contour interval of 0.5 mb). 42

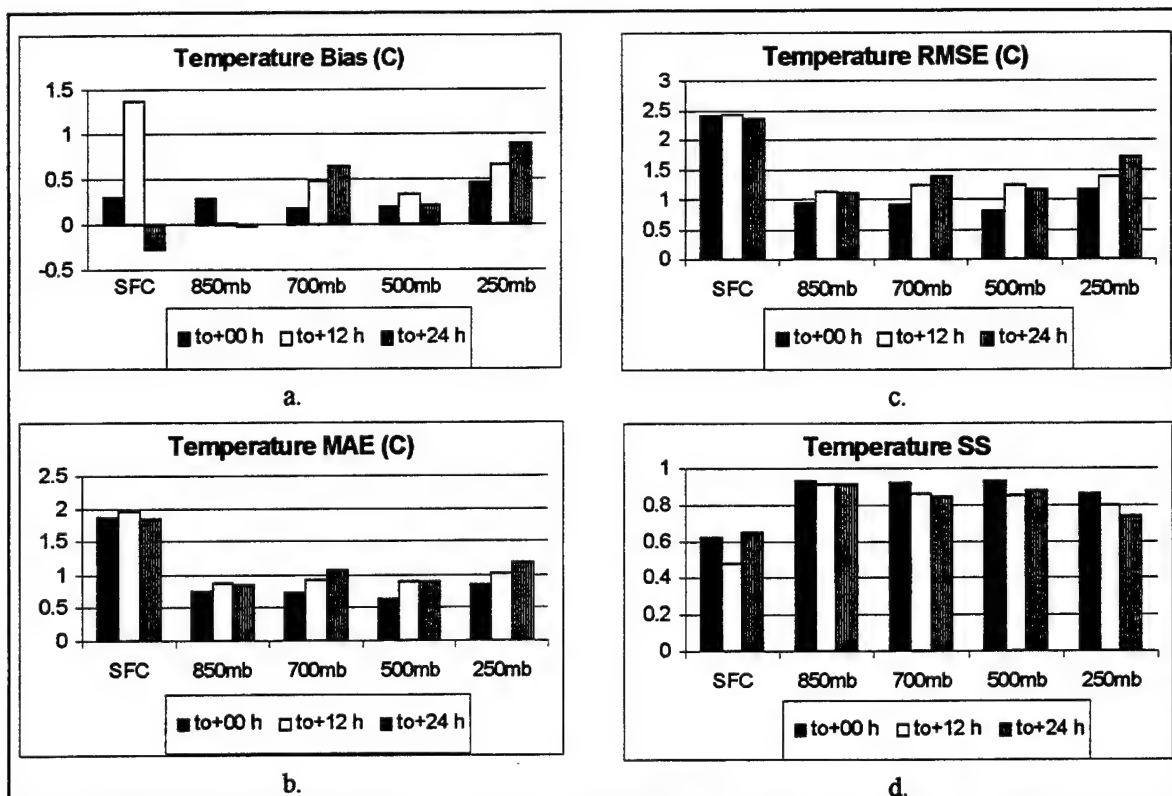


Figure 11. Fine-mesh temperature statistics for 1 June to 28 August 1997.

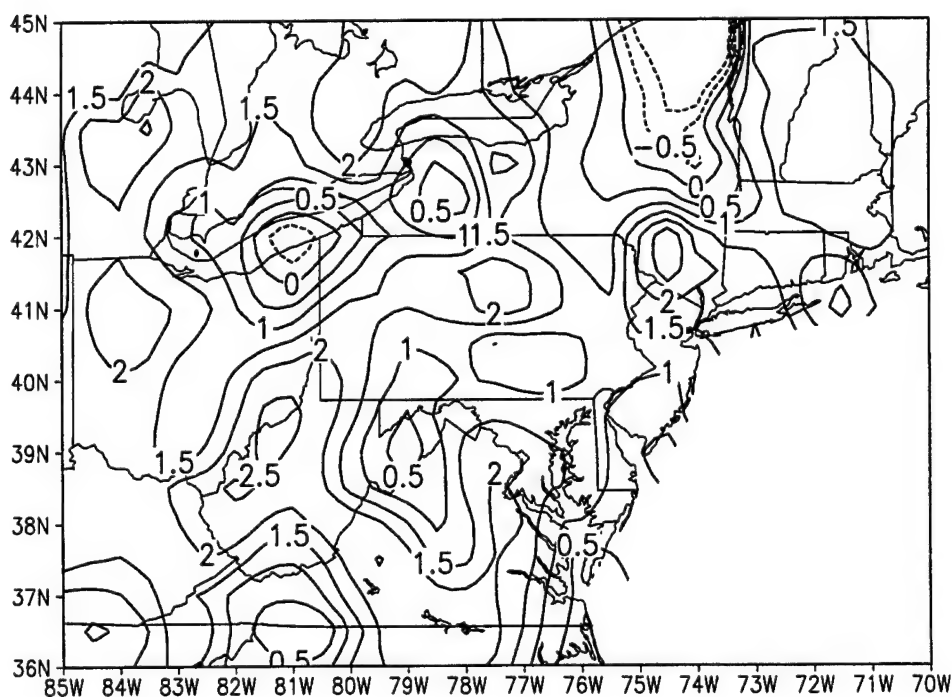


Figure 12. Fine-mesh surface-temperature bias, forecast time = $t_0 + 12$ h (contour interval of 0.5 °C).

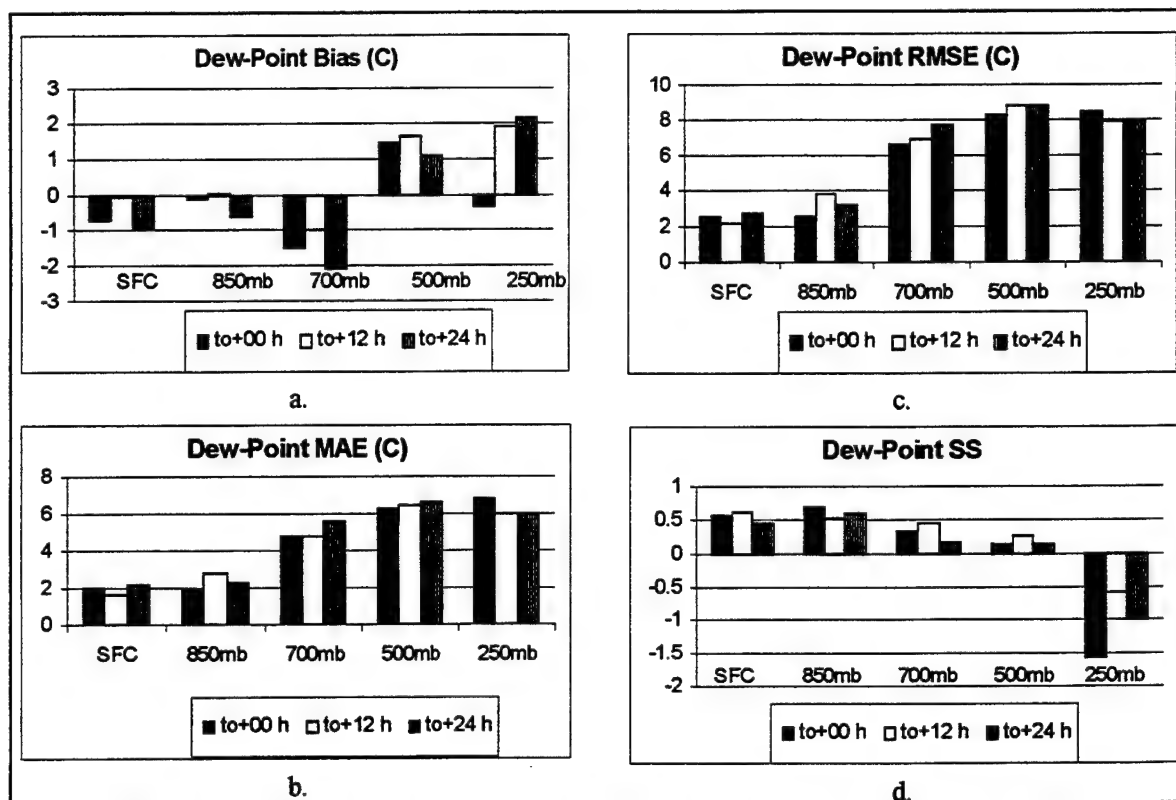


Figure 13. Fine-mesh dew-point statistics for 1 June to 28 August 1997.

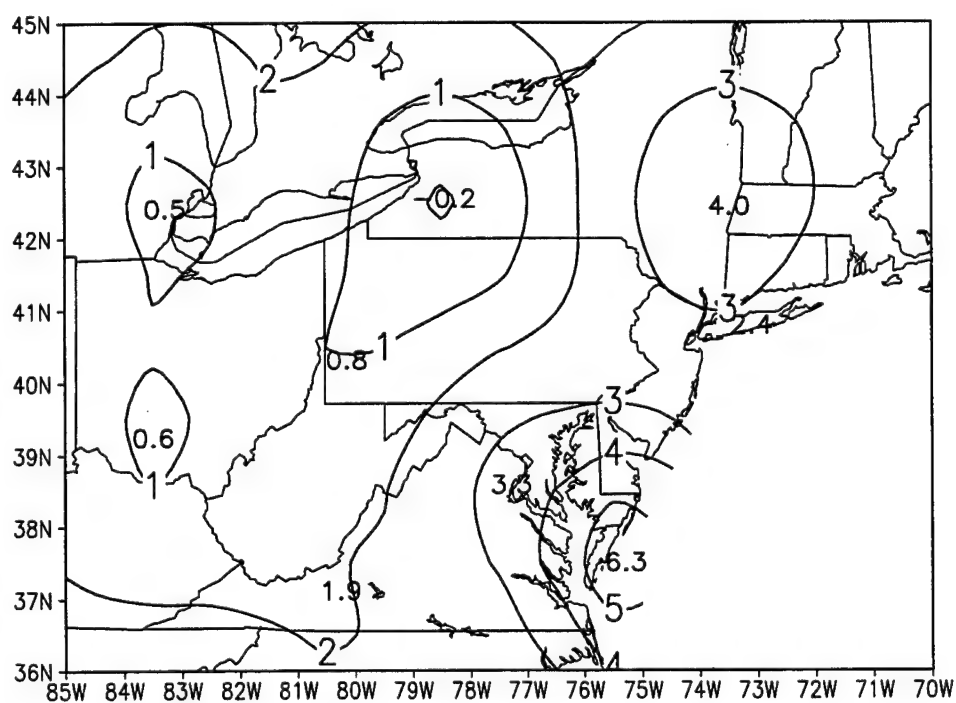


Figure 14. Fine-mesh dew-point bias at 250 mb, forecast time = $t_0 + 24$ h (contour interval of 1 °C).

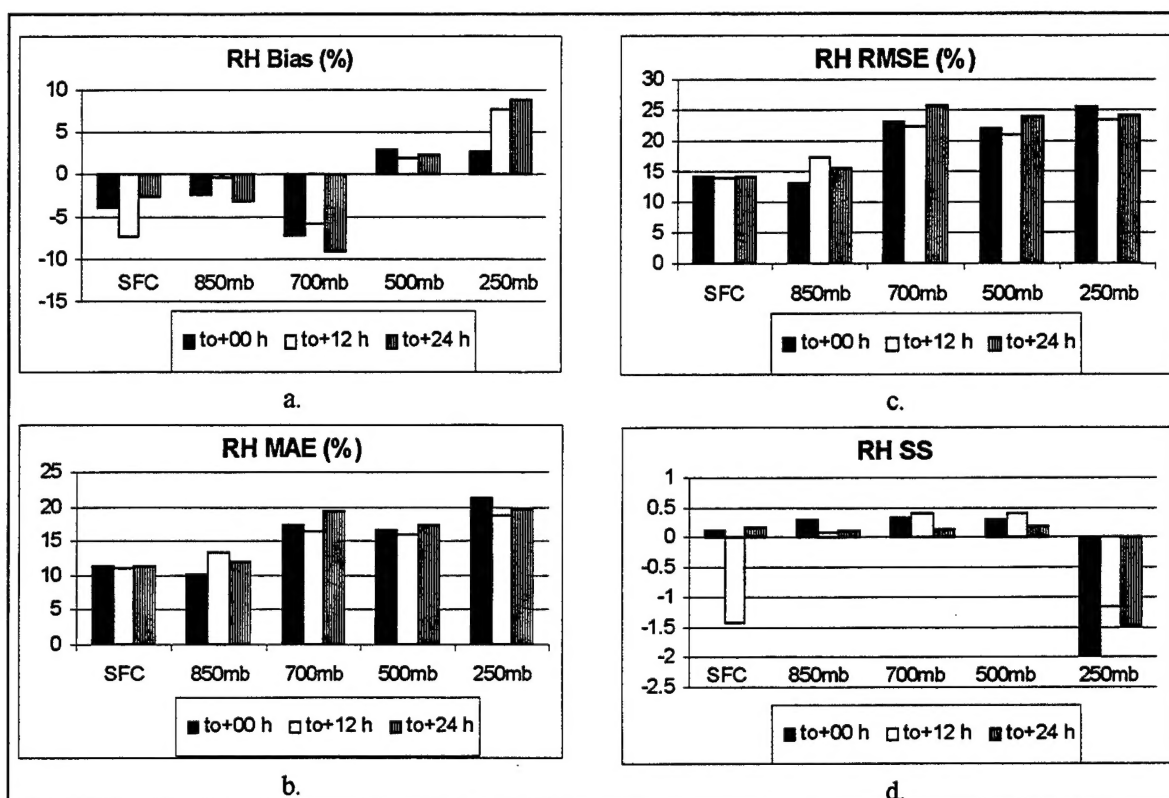


Figure 15. Fine-mesh relative-humidity statistics for 1 June to 28 August 1997.

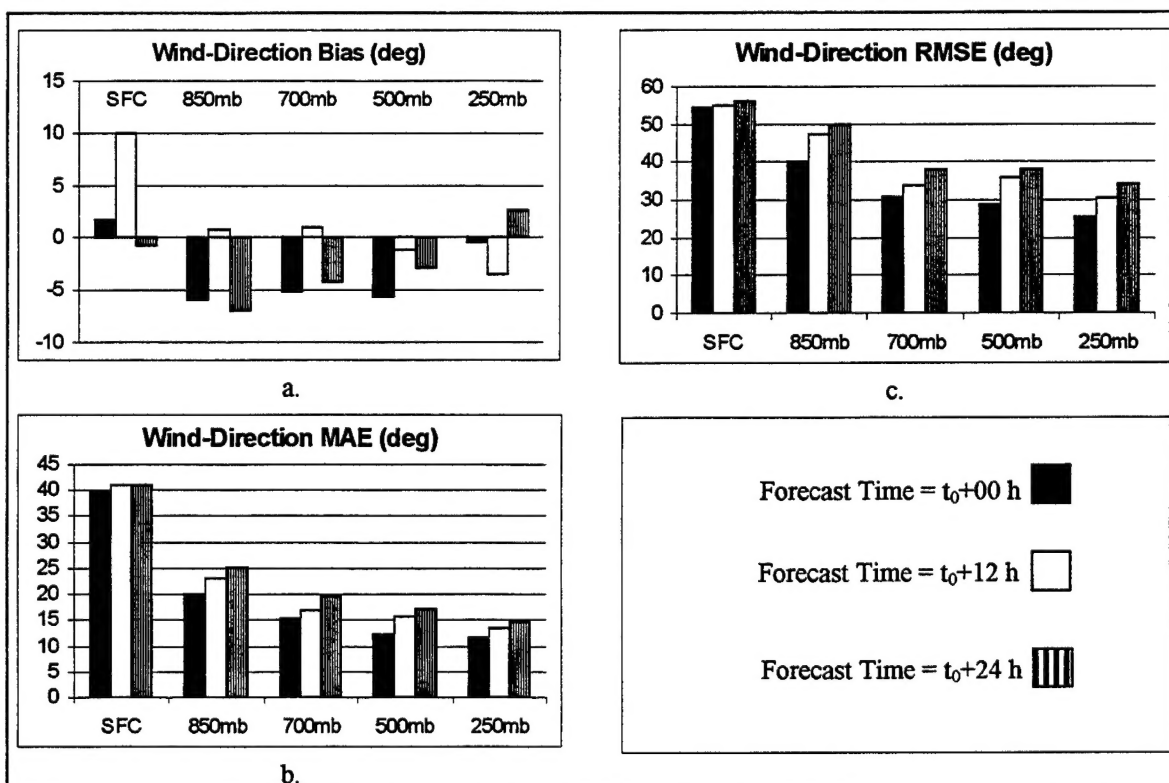


Figure 16. Fine-mesh wind-direction statistics for 1 June to 28 August 1997.

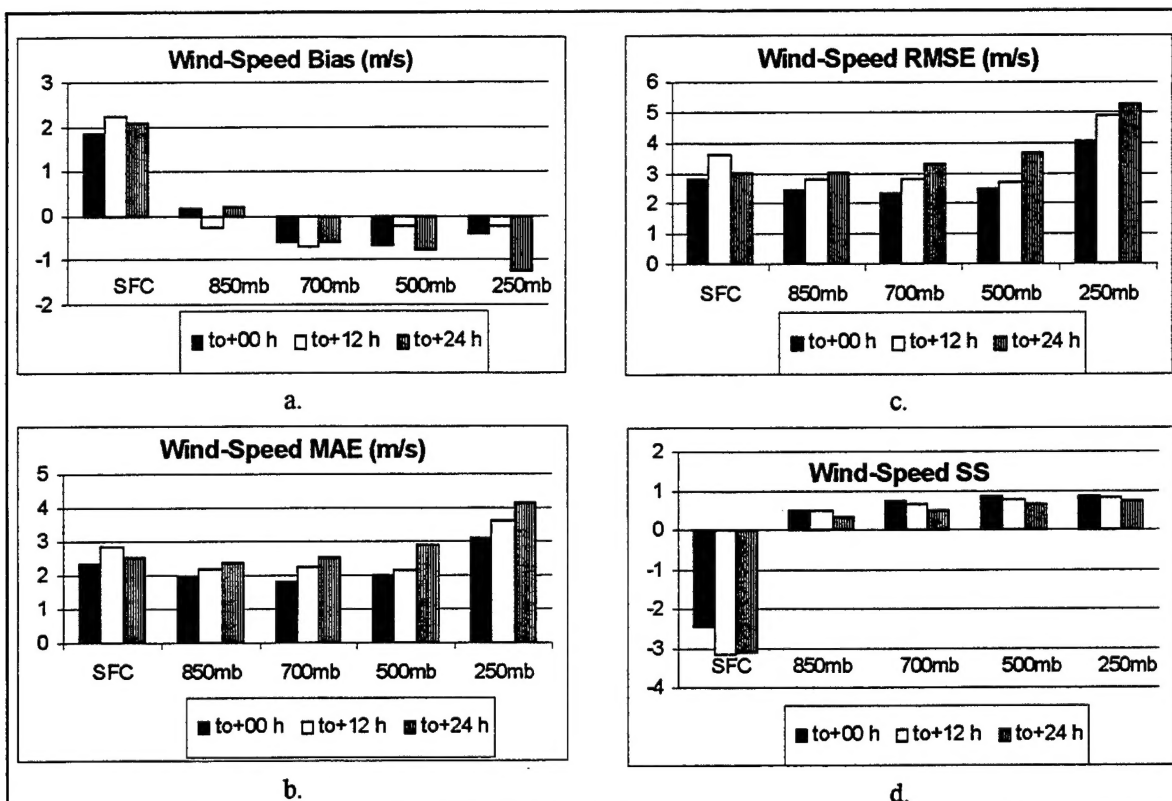


Figure 17. Fine-mesh wind-speed statistics for 1 June to 28 August 1997.

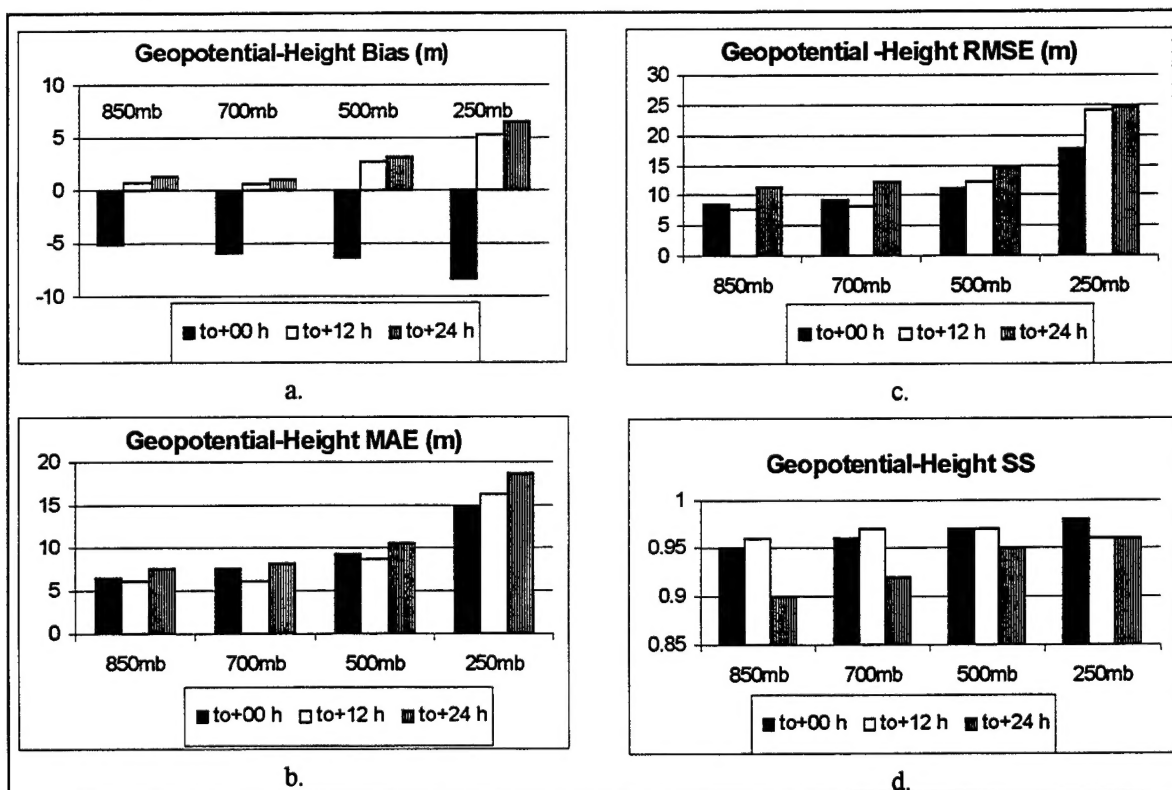


Figure 18. Fine-mesh geopotential-height statistics for 1 June to 28 August 1997.

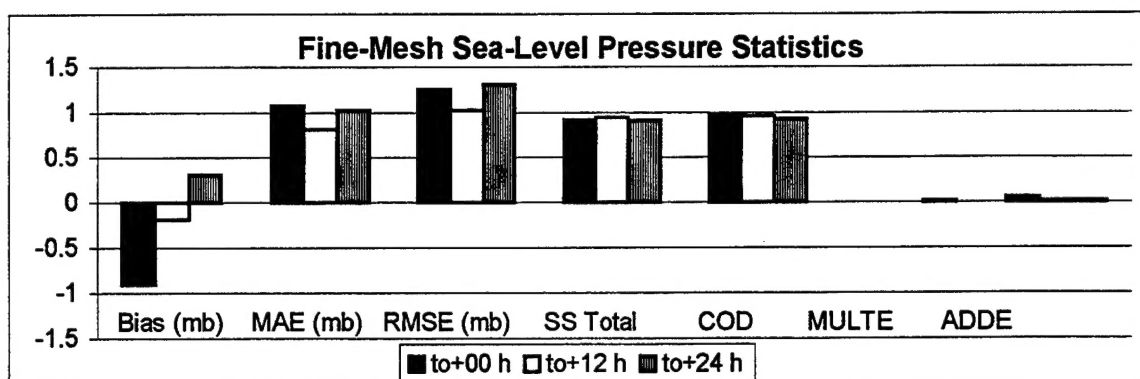


Figure 19. Fine-mesh sea-level pressure statistics for 1 June to 28 August 1997.

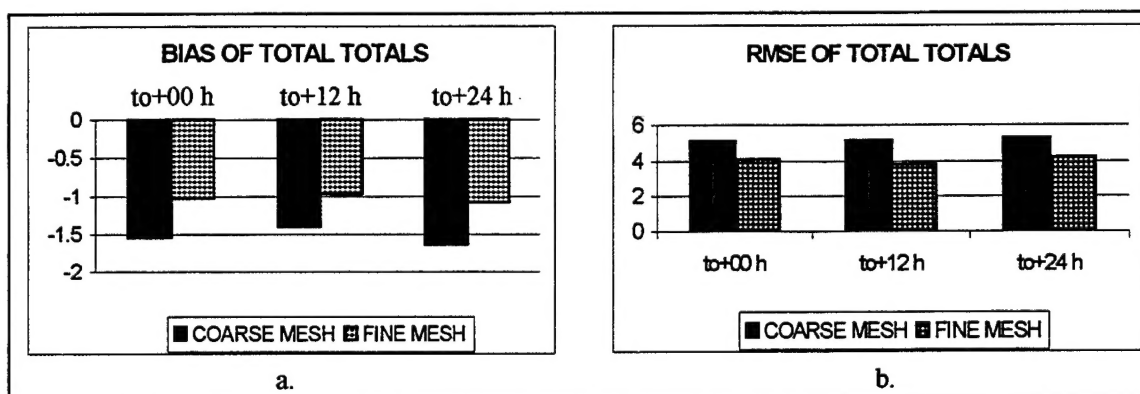


Figure 20. Total-totals index statistics for 1 June to 28 August 1997.

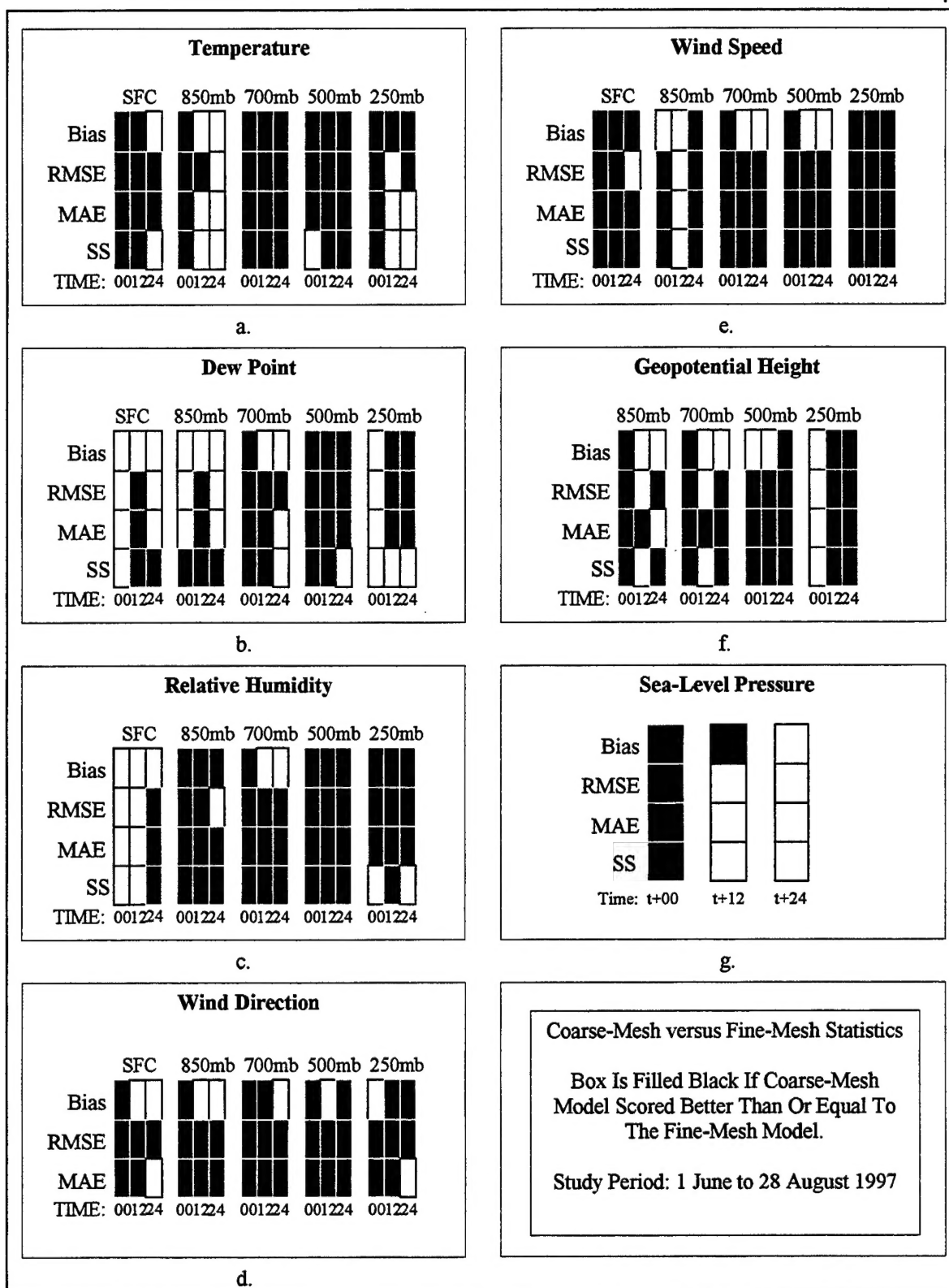


Figure 21. Coarse-mesh and fine-mesh statistical comparison.